

Clinton campaigns for health reform

US President Bill Clinton has made a brave start on his campaign promise to extend health care to those not now insured, but his plans will cost more than he predicts.

PRESIDENT Bill Clinton cannot be accused of failing to deliver on one of the most insistent themes of his election campaign a year ago. After months of closed meetings, punctuated by leaks to lobbyists and the press, Clinton has this week made public the essence of his proposal to reform the US health care industry. It is an achievement (for which the president's wife, Hillary Rodman Clinton, must take much of the credit) that last year's campaign skeleton has been hung with so much substantial flesh. Evidently the president and his colleagues in this enterprise have learned a lot. So have the major players in the health-care industry; 'socialized medicine', whatever that is, is not the only alternative to the free play of the health insurance market. But there is a long way to go before anybody knows exactly what kind of health-care system the Congress will allow Clinton to have.

Here, in essence, is how Clinton's first draft would work. Physicians, hospitals and insurance companies would voluntarily join together to form what is called a "managed care plan", offering a defined package of medical benefits for a fixed premium, either in a predetermined area or perhaps even nationwide. These will be the providers of health care. The federal government will say what basic medical benefits the package will include; individuals would be allowed to buy better coverage at their own cost.

At the same time, new entities, called "regional alliances", would be created under government supervision to represent people in various geographical areas. The alliances, essentially the purchasers of health care, would negotiate premiums with managed care plans in the area and then choose which plan to join. Premium payments would be handled centrally through the regional alliances and forwarded to insurance companies, which would either streamline paperwork or create a whole new layer of bureaucracy — depending on one's point of view.

Clinton's hope is that "managed competition" between managed health plans will save money. Certainly physicians and hospitals bound together in a plan would be under pressure from their partner insurance companies to justify the costs they incur. All parties to the arrangement would know that their plan might not win custom from the regional alliances if its costs for the basic health plan exceeded others on offer. That is the essence of managed competition. For what it is worth, it is a little like the arrangements recently introduced in the British National Health Service, where providers and purchasers have been sharply distinguished from each other. The rub is that there are few data in the

United States to show convincingly that current managed care plans, such as Health Maintenance Organizations in being for the past 20 years or more, actually provide high quality medicine at lower cost.

How will the uninsured benefit from these arrangements? Clinton's proposal hangs on the notion that access to health insurance should be coupled to employment, but not too tightly. To be sure, one of the more controversial aspects of the Clinton plan is a requirement that all employers pay insurance premiums for their employees, but with a ceiling of 7.8 per cent of payroll. But the poor or unemployed would still be eligible to purchase low-cost coverage through their regional alliances; those who could not afford to do so would be subsidized.

There are three crucial tests of whether Clinton's draft, and the iterations of it in the months ahead, will meet the need, but the first and most important test is whether it will indeed contain the growing cost of health care in the United States. What is called the health-care crisis is a real crisis. The United States already spends 14 per cent of its gross domestic product on health care, compared with 5 per cent three decades ago. The percentage is growing at 4.6 per cent a year in real terms. Industrialists (and managers of the federal budget) are alarmed that a continuation of this trend will consume what might otherwise be spent on industrial competitiveness or the reduction of the federal deficit. Many large companies, saddled with the requirement to pay generous workers' insurance premiums, find themselves at a disadvantage in the international marketplace.

The second test is whether the plan will meet the needs of the estimated 37 million people (many of them employed by small businesses) who have no insurance. It has always been anomalous that the richest country in the world should let many of its indigent and chronically sick fend for themselves, or look to chance and charity, when they need care. That does not mean that all of them are denied medical treatment. Indeed, their medical care at present costs roughly \$76 billion a year, paid for by the government through taxes (and programmes such as Medicaid) or by the insured, who subsidize the uninsured through higher charges than would otherwise be the case. But Clinton and the United States in general will benefit from the civility that accrues if the uninsured have a right to basic care. Clinton deserves credit for tackling this issue head-on.

The third test of the Clinton plan is whether it will survive politically. That boils down to the question whether man-



aged care will so improve the efficiency of the system that the United States can cut the rate at which costs increase and also provide universal coverage without raising taxes substantially. So little is known of the effectiveness of managed competition that there are only flimsy grounds for distinguishing that expectation from hope. Universal coverage will not come free. More seriously, if the plan as it stands were to be enacted intact, most middle-class Americans would either pay more for their current benefits package or lose some benefits they now enjoy. In a country in which the vast majority of people are used to choosing their own doctors and getting whatever services they want, that is asking a lot. But these are the rocks on which the Clinton plan is sure to be impaled in the months ahead. It will take all of a president's persuasive power to ensure it does not founder on them.

Nobody (probably not even the president himself) expects the Clinton plan to survive the US Congress in its present form. More than half-a-dozen congressional committees and more than a thousand special interest groups will have their say before any legislation materializes, at best a year or two from now. But some sensible reforms will emerge from the heated debate that will hold centre stage in US politics for quite a while. For that, Clinton deserves thanks.

Ending tribal wars

It is not too soon to plan how the rest of Europe will respond to the slow ending of the Bosnian fighting.

THERE are some positive things to say about the war in Bosnia and Herzegovina, now being dragged to an end by the exhaustion of the combatants rather than the success of the peace negotiators. For one thing, the war has not spread to Macedonia or to Kosovo. And Slovenia, the northernmost member of what was the Yugoslav confederation and the first to break away from it, has been uncannily silent during two years of fighting. Voluntary agencies and troops (under the aegis of the United Nations) have also shown great courage, even heroism, under trying circumstances. Otherwise, there is no good news. The war has been and remains a shaming business for participants and bystanders alike. It will cast a shadow over Europe for decades to come. All sections of the intellectual community therefore have an interest in knowing what went wrong.

The nature of the conflict, first in Croatia, then Bosnia, is central. Early on, the use of the term "tribal" to describe what was going on prompted a torrent of complaint from readers whose lives were being turned upside down. But even the passage of time does not suggest a more appropriate term. It hardly requires DNA fingerprinting of the population of ex-Yugoslavia to tell that this has not been an ethnic conflict. Nor, for want of homogeneity, has it been a religious war. To be sure, religious Serbs are mostly Orthodox Christians, church-going Croats are mostly Roman Catholics, but the Muslims of Bosnia are indistinguishable from those who

used to live next door to them by the variety of slavonic language they speak.

The wellsprings of mutual hostility are rather to be found in the more familiar cultural attributes that tie some groups together, dividing them from others. Shared traditions (real or imagined), family ties and rural people's attachment to the land their forebears worked are both cementing and divisive; language and religion then confirm people's sense of kinship with members of their own group and their sense of being at odds with other groups. The past two years have horrified the rest of Europe and the world because of the general supposition that European civility is now powerful enough to damp down ancient antagonisms and, in particular, to prevent people killing other people simply on account of their group membership. Events, to everybody's dismay, have shown that not to be the case.

What is to be done? Mr George Soros, the financier-philanthropist (who is spending \$50 million for relief in Bosnia, not to mention \$100 million for assistance to research in the ex-Soviet Union) argues in a privately published pamphlet that the rest of the world's error over Yugoslavia is its willingness to accept the creation of numerous independent states whose basis is essentially ideological. He argues that military intervention in the crumbling Yugoslavia would have been justified to prevent such an outcome. The belief that ideologically motivated states tend to be illiberal is easily supported, but many are already so founded. There is, for example, the Islamic Republic of Pakistan, not to mention Iran and the Sudan (or at least the northern part of it). These anomalies will not melt away simply because the rest of Europe is ashamed of what has happened in Bosnia.

The more immediate question is that of how the rest of Europe should deal with what is left of Yugoslavia when the war eventually drags to an end. It is a haunting question: revulsion at recent atrocities will blunt compassion for communities uprooted from their homes, robbed of breadwinners and turned into refugees. In Bosnia, all three communities have been responsible for acts of barbarism. But the rest of Europe's interest is that the past two years should reinforce as little as possible the tribal hostility, which argues for generous assistance with rehabilitation coupled with plain statements of the reasons why the illiberality of the recent war is offensive. The second ingredient is as important as the first.

A report on page 287 of this issue describes what Germany is already doing to rehabilitate research in Croatia. The motives are admirable: to focus research on the needs of Croatian industry and thus to strengthen the local economy. The open question is whether this assistance is tempered with the essential message that the tribal wars of the past two years themselves offend against the principle that science is an international enterprise. The University of Zagreb (the capital city of Croatia) used to be a polyglot university, for example. Will it be so now, and in the future? Germany could exert a beneficent influence on that important question, which will recur, as recent events in the Caucasus have shown.

Russian science gets caught up in struggles over political reform

Moscow. The Russian president, Boris Yeltsin, has written to the Minister of Science, Boris Saltykov, and the president of the Russian Academy of Sciences (RAS), Yuri Osipov, demanding proposals from them for a radical reorganization in the way that Russian science is organized. They have been given a month to reply.

In particular, Yeltsin has asked for suggestions about ways of funding individual scientists on a competitive basis. "The country is not able to finance non-productive scientific research", said Yeltsin in his letter. "It is necessary to find new approaches to financing scientific research within a market economy, and to supporting those scientific centres that are concentrating on the key trends in science."

Yeltsin's appeal coincided with a meeting between a group of scientists and the head of the Russian parliament, Ruslan Khasbulatov, at which Khasbulatov — one of Yeltsin's chief political rivals — also spoke of the need for a recovery of science in Russia. He blamed Yeltsin and the government for the current crisis in science, and promised the scientists that the parliament would give them help.

Both politicians appear to have been acting pragmatically. The conflict between them has reached deadlock, and as a result, the two branches of the Russian government are each trying to attract as many supporters as possible.

This situation has made the scientific community an attractive target. The evaporation of research funds, the temporary closure of many research institutes, and a failure to pay salaries that are already pitifully small have persuaded many scientists to reconsider their previous support of the government, and to look closely at the difficulties of bringing democratic reforms to life. In turn, the politicians now trying to carry through these reforms are worried about losing the backing of the Russian scientists, who have until now been among their most active supporters.

Yeltsin appears genuine in his desire to preserve the scientific potential that still remains in Russia. And the text of his letter reflects suggestions that scientists themselves put forward at a conference organized two years ago by the academy.

At the time, their suggestions were rejected by the academy, which continued to turn for support to the structures and procedures that had previously existed for financing scientific research in the Soviet Union.

As yet, there has been no official re-

sponse to Yeltsin's letter from either the academy or the ministry. However, according to one informed source, its contents have been discussed at a closed meeting of the presidium of the academy, where it was interpreted as a direct attack.

The letter has already aggravated tension between the academy and the Ministry of Science. The two were already in open conflict after the dismissal of Andrei Gonchar, vice-president of the academy,

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Rivals Yeltsin and Khasbulatov, speaker of the Russian parliament, vying for the support of the scientific community

from his position as president of the Russian Foundation of Basic Research.

The academy's academic secretary, Igor Makarov, has claimed that the Ministry of Science was responsible for drafting Yeltsin's letter, with the intention of destroying the system of academy institutes. The academy's president, Yuri Osipov, has reacted more calmly; he told members of the presidium that, as far as he was aware, the ministry was not involved, and the letter was prepared at the initiative of one of the president's advisers.

Whatever its origins, the Ministry of Science has, in contrast to the academy, reacted to the letter with satisfaction. "We have been following for a long time the scientific and technical policy now being demanded by the President of Russia," says Andrei Fonotov, the First Deputy

Minister of Science.

Fonotov points out that the Russian Foundation of Basic Research was created at the initiative of the ministry, as was a system of scientific centres. "We first stated long ago the need of making serious organizational changes within the management of Russian science. There is nothing that we will have to correct in connection with the President's letter."

In addition, the ministry recently announced plans for creating a new institute of "state professors", a suggestion made by a number of Russian academicians. A scientist awarded such status will receive a stipend of about \$1,000 a month — a large amount by Russian standards — and other benefits are likely to be introduced.

The status of a state professor will be awarded only to scientists who have been recognized as "outstanding". The ministry estimates that up to 5,000 such appointments will be made. This implies the creation of another state foundation with a yearly budget of US\$50–\$60 million.

The effect of Yeltsin's letter remains unclear. However sensible the scientific and technical policy followed by the ministry, it is unlikely to prevent the rapid collapse of Russian science if the government continues its financing in the same meagre and irregular manner as it has done this summer.

The Ministry of Economics has now paid money whose absence caused a storm of outrage in July in the scientific community. But it not yet come up with any money towards the scientific budget for the second half of the year.

Vladimir Pokrovsky

■ As *Nature* went to press, we learnt that President Yeltsin has issued a decree giving urgent help to Russian science. Details will be published in next week's issue.

Concern raised over Warren Spring move

London. British scientists opposing the plans by the Department of Trade and Industry to merge its Warren Spring Laboratory, one of the country's main centres for research into environmental technology, with AGA Technology, have received backing for their concerns about the government's action from the Royal Commission on Environmental Pollution.

In a letter sent to the department last week, Sir John Houghton, chairman of the commission and a former head of the Meteorological Office, asks for reassurance from

the government that the merger of the laboratories will create a national organization "which can quickly achieve a high reputation for expertise and impartiality."

Sir John also emphasizes that the new joint organization, to be known as the National Environmental Technology Centre, should not only take over key staff from Warren Spring but retain and develop some of that laboratory's key characteristics. The staff have claimed that the high number of scientists reluctant to move will make this difficult to achieve. □

UK scientists threatened with pay freeze

London. Thousands of British scientists working in universities and government-funded laboratories face the prospect of a pay freeze next year as a result of moves by Kenneth Clarke, the chancellor of the exchequer, to reduce public sector spending.

Clarke announced last week that the government intends to keep the overall salary bill for government employees for the financial year beginning next April to the same level as this year. He added that increases could be paid where there was evidence of productivity gains.

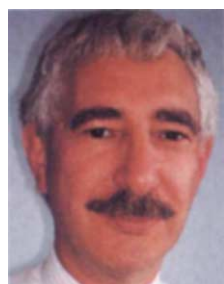
But organizations representing scientists in their salary negotiations with the government point out that this formula means that increases can be paid only where there is a net reduction in staff.

They also claim that a freeze on scientists' salaries conflicts directly with the message of the recent white paper, which expressed the government's concern about the state of British science — including its low appeal to school students — and its determination to give science a more attractive image. "The government says it is worried about people not going into science, but they seem to be doing the best they can to put them off even more", says Valerie Ellis, assistant general secretary of the Institute of Professionals Managers and Specialists (IPMS), whose members include scientists from many government laboratories.

Scientists have already had salary increases held back by the government. This

year — like all public employees — they have been limited to a rise of 1.5 per cent, considerably less than the anticipated rate of inflation for the year. If Clarke sticks to his word, next year's situation will be worse.

Particularly hard hit, according to Ellis, will be those in certain grades (such as chief scientific officers) who were to have had a salary review next year to bring them up to comparable salaries in the private sector.



AUT president David Triesman warns of a crisis in universities.

Such reviews are now likely to be postponed.

University staff have reacted scathingly to Clarke's promise of salary increases to those whose productivity has gone up. The Association of University Teachers (AUT) points out that, as a result of a rapid rise in student numbers, staff in the "old" universities are today teaching 46 per cent more students than they were six years ago; but their salaries have actually decreased by one per cent in real terms over the same period.

AUT officials — in common with the Committee of Vice Chancellors and Principals — also argue that the decision to continue to hold down salaries comes at a particularly sensitive time. Many university

teachers are rapidly approaching retirement, but faculties lack the funds to make the new appointments needed to ensure an orderly succession of good academic staff.

"Taken together with the continuing fall in morale among researchers, the chancellor's announcement could produce one of the most acute crises that we have seen in British universities this century", says AUT president David Triesman.

Officially, Britain's research councils say they are withholding comment until the formal announcement of the terms for next year's pay deal with public sector employees, which will be unveiled in the autumn budget statement at the end of November. Behind the scenes, science administrators are also hoping to put pressure on Whitehall (and sympathetic ministers) to convince the Treasury that a salary freeze on scientists would be counter-productive in the long-term. Meanwhile, the IPMS is planning a series of demonstrations later in the autumn to take their case to the public.

David Dickson

UK graduates a bridge to Japan?

London. British companies should consider sending their graduate recruits to Japan for a year before they start work in Britain. The suggestion comes from the chairman of a committee set up by the Office of Science and Technology (OST) to advise the government on how to strengthen its scientific and technological links with Japan.

Sir Geoffrey Allen, a former chairman of the Science and Engineering Research Council, and now an adviser to the Japanese company Kobe Steel, said last week that such a move would help build a greater awareness in British industry of the reasons for Japan's industrial success.

It would also avoid disputes over the protection of intellectual property that frequently plague such arrangements. At present, managers are often reluctant to send their scientists and engineers to work with companies in Japan because of the knowledge they inevitably take with them.

Sir Geoffrey was speaking shortly before this weekend's departure of William Waldegrave, the science minister, for a five day visit to Japan. One of the main purposes of the visit will be to encourage more Japanese companies to set up research and development facilities in the UK.

The new advisory committee is made up of ten academics and industrialists, each with personal experience of Japan. According to OST officials, the committee will be responsible for analysing and making recommendations on the strategic issues concerning the development of closer scientific and technological links. □

Alison Abbott

Germany to boost R&D in small companies

Berlin. Germany's ministry of research and technology (BMFT) has launched a new programme designed to boost research and development in small and medium sized companies employing up to 500 employees.

Particular emphasis will be placed on companies in the new *Länder* of what was formerly East Germany. But a report from Berlin's Centre for Social Science Research (CSSR) criticizes previous programmes run by the BMFT with similar goals, claiming that their success has been limited by the general underfinancing of the *Länder's* industrial base.

The new BMFT programme runs until 1997, and will have a total budget of DM200 million (\$320 million). The money will be available to support collaborative projects between two or more companies, or between a company and a research institute. It can also be used to enable company employees to spend short periods working in a research institute.

East German companies will be expected to match the money put up by the ministry with their own funds. Those from the old *Länder* will have to provide more — be-

tween 60 and 75 per cent — of the project money themselves.

The CSSR has conducted assessments of many government projects, particularly in the new *Länder*. In a report issued last month, it warns that previous programmes intended to bolster industrial R&D in the former DDR have in fact done little to stop its steady decline.

At the request of the BMFT, the centre investigated the results of two special programmes funded by the ministry which came to an end last year. One, costing DM50 million so far, is aimed at increasing the number of research staff in industrial companies. The other (at DM62 million) to help companies develop their ability to carry out contract research.

Both programmes were surprisingly successful in their immediate aims, says the centre. But despite this, it adds, the eastern German companies often did not have the capital to move new products through development, effectively wasting the research effort. And overall industrial research capacity continues to fall.

Brandenburg's new university feels the pace of change

Potsdam. The newly founded University of Potsdam in the east German state of Brandenburg will finally elect its first independent senate by the end of spring, according to a statement issued last week by Brandenburg's ministry of science, research and culture.

The university will have been built virtually from scratch in less than three years. But the pace of the university's establishment has still been criticized as too slow. Delays have created an uncertain working atmosphere for new staff, and stress on temporary staff, who have not yet been told whether they have a future at the university.

After reunification in October 1990, Germany's science advisory council, the Wissenschaftsrat, reviewed the science and higher education infrastructure in the former East Germany. In Brandenburg this procedure was complete in the spring of 1991.

The Wissenschaftsrat recommended the founding of three universities, the largest to be in Potsdam, the state capital, and two others at Cottbus and Frankfurt-Oder. The federal government agreed to pay nearly half of the DM3 billion (US\$1.9 billion) cost of building the new establishments.

The Brandenburg government reacted very quickly, passing the first university act in the new *Länder* in June 1991. The act anticipates a total of 34,000 student places by the year 2000. Around 60 per cent of these will go to the three universities, with 10,500 at Potsdam, and 40 per cent to the five newly founded, technologically orientated *Fachhochschulen*.

Potsdam University is being developed at the site of the Brandenburgische Landeshochschule. This was established in 1948 with the intention of becoming a university. But when East Germany was founded in 1951, the status of the institute was reduced to that of a *Pädagogische Hochschule* (teacher training college).

After the Berlin Wall came down in October 1988, the *Hochschule* readopted its original name. At that time it had 120 professors and more than 500 other academic staff. In September 1990 it held its first free elections for a senate, and in July 1991 the college was re-established as a university.

A non-elected temporary founding senate was put in place by the ministry of science. Originally intended to sit for only

two years, the founding senate found that the process of setting up the university's structure took much longer than expected. Brandenburg's science minister now says that the first independent senate elections will take place "during the winter semester". But the continuing delays have caused some unhappiness, particularly among the former academic staff of the *Hochschule*.

Reunification put enormous pressure on these staff. They had to be evaluated for their "personal and political integrity", as well as their professional competence to do their work. The evaluation committees, which consisted of professors from the new and old *Länder*, and academic and student representatives, completed the task by the beginning of 1992.

Twenty-five of the 120 professors were sacked for collaboration with the secret police, and a further fifty left of their own accord. The remaining professors were recommended as being suitably qualified to continue their jobs.

Even then, however, there was no immediate peace for them. The founding senate set up a second round of evaluations, this time with more specialist committees. This process is not yet complete. But 30 have so far been positively evaluated, and their positions are waiting for approval by the ministry of science. A further seven have

not made the grade, and they will either be sacked or offered lower positions.

The fate of more than 500 other nonprofessorial academic staff (the *mittelbau*) remains uncertain. Around 90 per cent were positively evaluated by the Wissenschaftsrat and put on interim contracts while the new structure was being developed. But the new university will not be able to employ all of them and about a hundred are likely to lose their jobs.

Like all establishments in east Germany, the university has the right until the end of this year to dismiss staff if they are surplus to the requirements of the new structure. But procedures are moving so slowly that this deadline may not be met.

Dismissals after the deadline will be much more complicated because, in the former East Germany, nearly all university academics were on permanent contracts. Unless this is changed before January — the university wants to offer short-term contracts to 80 per cent of the *mittelbau* it retains — they will be guaranteed a job for life because academic staff are civil servants in the federal republic.

The *mittelbau* will not be whittled down in all departments. Much of the science faculty, for example, is being built up. The former biology department had five professors and 35 *mittelbau*; it is now to have 17 professors and will increase other academic positions accordingly. Several entirely new 'interdisciplinary' departments such as nonlinear dynamics and cognition, are being created — an enormous scientific advantage for the university, according to Jürgen Kurths, a chaos researcher.

Potsdam University is keen to establish a good social mix among its professors, and a good spectrum of ages. According to Gerhard Kempter, a chemist and deputy rector of the founding senate, around a quarter are likely to be elected from the new *Länder*. The rest will be Westerners.

Alison Abbott



University of Potsdam

AIDS foundation plans three centres

Paris. The opera singer Luciano Pavarotti has pledged part of the proceeds of an open-air concert in Paris last month — and those of future concerts — to the AIDS foundation set up earlier this year by Luc Montagnier of the Pasteur Institute and Federico Mayor, director-general of the United Nations Educational Scientific and Cultural Organization (UNESCO).

Donations from Italian banks and individuals have already provided the "World Foundation for Research and Prevention of AIDS" with around US\$650,000. These donations have covered the foundation's start-up costs. They have also covered the first phase of its prevention and social programmes in Africa, including a FF5 million (\$900,000) project to educate children whose parents have died of AIDS.

The foundation's funds are still far short of the \$30 million or so it will need to finance three planned clinical research centres. But Pierluigi Vagliani, an adviser to Mayor, says he is confident that income will rise sharply in the new year.

The Saint Joseph Hospital in Paris has provided the foundation with a three-storey building for the first centre. The government of the Ivory Coast has supplied land for a second. And Montagnier is negotiating with an institute in California to build a third, says Vagliani.

Each of the three centres will concentrate on Montagnier's research into cofactors (see *Nature* 361, 287; 1993). But Vagliani says that the foundation is actively seeking international participation to open other new areas of research.

Declan Butler

Canada's Liberals promise money for science

Quebec. Canada's Liberal Party has promised to double the amount of money the country spends on research and development if it wins the federal election on 25 October. The party leader, Jean Chrétien, says that a Liberal government would invest an additional C\$1 billion in R&D over the next four years.

The Liberal party, which is challenging the ruling Conservatives, claims that the country's principal failure in dealing with research and innovation lies in its inability "to gather, finance, administer, diffuse and market technology".

As a result, it promises to establish a "technology network" linking universities, industry associations and both federal and provincial governments, in order to make scientific and technical information more

readily available to industry.

But it also says that its pledge to double R&D spending will depend on the country "demonstrating its ability to absorb and manage such an increase".

The promised increase in the research budget forms part of a package of promises totalling an estimated C\$5.35 billion over four years. The extra costs would be offset by cutting C\$7.1 billion from existing public programmes.

A large part of the reduction would come from cancellation of the C\$5.8 billion, 13-year programme to purchase military helicopters. The Prime Minister, Mrs Kim Campbell, when she was defence minister, described the programme as untouchable. But once an election was under way she agreed to reduce the number of helicopters

to 43 from 50. Chrétien says he now wants to eliminate all of them.

The Liberal party's promise of support for science comes at a time when the ruling Conservatives have been severely criticized for their poor record in science funding. Perhaps the most serious blow to the Conservatives' credibility has been the resignation of Pierre Perron as president of the National Research Council (NRC) almost a year before his five-year term of office was due to end.

In a letter to the council's employees at the end of last month, Perron outlined the serious implications of deep budget cuts and what he described as the "anaemic support of R&D" by the government. The latest cuts, announced at the end of July, will raise the potential annual shortfall to some 10–15 per cent of NRC's operating budget between 1995 and 2000, Perron said.

"If Canada is to catch up with the average of the G-7 countries, governments at all levels would have to more than double their investments in R&D over the coming years, while industry would have to increase theirs by a factor of three to four", he wrote. "Some 15,000 to 20,000 new jobs in R&D would have to be created every single year for almost a full decade."

Perron claimed that the government was overwhelmingly preoccupied with "the rate of growth and the national debt". As a result, the focus is on potential short-term savings, ignoring negative long-term impacts.

For this reason, NRC has decided to present a case for its next long-range plan (1995–2000) early in the fall, Perron said. It will seek a new funding base of some C\$260 million a year for its operating budget (an increase of C\$10 million), and additional one-time funding of C\$112 million for R&D in key areas.

David Spurgeon

Information superhighway takes off

Washington. The Clinton administration has published a nine-point "agenda for action" designed to encourage the construction of a national information infrastructure which, it hopes, will provide a high-speed telecommunications link to every home and office in the United States.

Vice President Al Gore said last week that the federal government would act as a "catalyst" for the infrastructure, which will be built by the private sector. "No one should have any doubt how important this is to this administration", said Gore.

The agenda published by the White House promises to ensure that the so-called information superhighway is available to all "at affordable prices", that the intellectual property rights of its users and service suppliers will be safeguarded and that the superhighway is compatible with international networks.

Commerce secretary Ron Brown says that his department has already discussed the laws needed to establish the network with 300 interested parties, one hundred of which were involved in drafting the 24-page agenda. The document promises to formalize consultation by setting up a 25-member advisory council.

Brown has invited nominations for the council, to be announced by December. Its make-up has not yet been settled. But universities and schools that want to use the network are likely to be represented. So, too, are industrial and commercial users, as well as the telecommunications and computer industries that will build it.

Legislation authorizing government agencies to provide several hundred million dollars of research and development funding for the proposed infrastructure is going well, and may even be passed next month.

The research and development law has already been passed in the House of Representatives, and an equivalent measure is being considered in the Senate as part of the larger National Competitiveness Act. Once the Senate act is passed, it will have to be reconciled with two corresponding Acts already passed by the House (see *Nature* 363, 104; 1993).

The parts of those laws dealing with the information infrastructure are broadly consistent with each other. They call for federal agencies such as the National Science Foundation and the National Institute of Standards and Technology to develop applications for the superhighway and to set up prototypes which will be transferred to the private sector as quickly as is practicable.

Colin Macilwain

Centre will co-ordinate Arctic research

Helsinki. In a move to coordinate the efforts of research organizations and initiatives studying environmental change in the Arctic, the International Arctic Science Committee has set up a Global Change Programme Office, based at the University of Lapland's Arctic Centre in Rovaniemi, Finland.

The office is headed by the director of the Arctic Centre, Manfred Lange. It will support and help to implement scientific projects concerning climatic changes and their effects on the Arctic region. According to Lange, one of its main functions will be to help to avoid overlap between the various programmes which address Arctic issues.

The creation of the new office is also intended to help focus on the pivotal role of the Arctic in climate change, covering both

the impact of change on the Arctic and the resulting feedback. "Given the fragile balances that sustain life in the Arctic and Antarctica, a rise in the temperature of 4 to 6 degrees over about a hundred years, as predicted by some computer models, might have devastating effects on polar ecosystems", says Lange.

The new office in Rovaniemi is intended as an information clearing house, linked to the multidisciplinary activities of the Arctic Centre. In addition to monitoring and advising on projects in the natural sciences, the office aims to bridge some of the gaps between these and Arctic research in social and economic fields, thus widening the scope of scientific enquiry to cover human influences on the natural environment.

Mark Waller

Croatian science looks ahead as it fights for survival

Zagreb & Osijek. Despite the frequent bombardment of laboratories, an insignificant research budget, and a continuing brain drain, Croatian science remains alive. Freed of the communist regime, the government is now planning to establish a new, competitive research system; and Germany is offering to help.

At present, survival rather than strategy is the name of the game. The war in the former Yugoslavia has created a deep economic crisis. Croatia's investment in science has dropped sixfold over the past three years. The \$30 million spent on research in 1992 — the same as the monthly sum spent supporting its 500,000 refugees — barely covers salaries and journal costs.

The average pay for a scientist has fallen to around \$180 a month. Croatia's four universities have given up hope of maintaining modern facilities; the department of chemistry at the University of Zagreb last year had only \$30 per student to spend on chemicals, software and equipment.

The problems facing Croatia's 9,000 or so scientists are not limited to the lack of money. Fighting has caused extensive damage to many research facilities. Worst hit is the University of Osijek in eastern Croatia, founded in 1975. Situated only 800 metres from the front line, its faculty of agriculture and Institute for Food Technology have been under fire for months, and several entire buildings have been destroyed.



Research minister Jeren is planning reforms

Staff and students have been forced to find premises elsewhere in the city. Research has inevitably been badly affected. But scientists have managed to repair some pieces of equipment, and are trying to continue their work in makeshift laboratories. "We didn't stop teaching for a single day", says the university's rector, Stanislav Marijanovic.

So far, the University of Zagreb, which, with 42,000 students, is Croatia's largest university, and is also distant from the front line, has escaped serious damage. But it is suffering from the brain drain that has affected the whole country, with an estimated 1,800 Croatian scientists and engineers now working abroad.

Research in many fields is still proceeding, often helped by foreign collaborators.

But Sibela Jelaska, a professor of molecular biology at Zagreb's Rudjer Boskovic Institute, complains that young people no longer consider doing a PhD in Croatia. "Why should they?" she asks. "We don't even have equipment for sequencing DNA in our labs."

As a result of the war, much international financial support for research has been withdrawn. The European Communities (EC) have frozen Croatia's participation in the PHARE programme, which finances the restructuring of post-communist countries such as Poland and Hungary.

According to Jean Luis Blanc, an EC official, it is Croatia's involvement in the war that has ruled out any research support from the EC. Croatia's science and higher education minister Branco Jeren sees it differently. He claims that opportunities for collaboration have been unfairly prejudiced by international sanctions being enforced against Serbia.

Others claim that restrictions on helping Croatia are now justified by Croatia's own conduct in the war. But Jeren is looking to the future. In particular, he is preparing a new research system to replace that inherited from the communist regime when it handed over power in 1990.

Before that, Yugoslavia had lacked sufficient funding to establish modern research facilities, and had no coordinated science policy. Few formal procedures existed for linking researchers and industry. Nor was there a system for creating links between university departments, or with researchers from other countries.

Jeren, a physicist who has been minister since February, is now planning substantial changes. Two new laws are expected to be approved by parliament later this month. One, a Science Act, will enable research institutes to shift from self-management to a more efficient and competitive system of administration.

The second, a Higher Education Act, will create similar changes in universities, allowing them to establish management hierarchies and to place real power in the hands of deans and rectors. The act will also limit student places, and allow universities to select the most talented students by raising entry requirements. Jeren says he is keen to support Croatia's young scientists by allocating 12 per cent of his research budget to a "new blood programme".

Many foreign countries, including some within the EC, have been reluctant to offer anything more than humanitarian aid to Croatia because of its conduct in the war. Several, however, are continuing to provide assistance to individual scientists. And Ger-



Bomb damage at the University of Osijek

many has now gone further by offering direct support for Croatia's efforts to restructure its science.

Germany's role has historical roots. As a country with traditional links to the region — as well as a desire to increase its influence in both cultural and educational matters — Germany was the first to recognize Croatia as an independent state in 1991.

It is partly motivated by direct self-interest. Germany is keen to stem the flood of Croatian refugees by helping to secure a stronger economy in Croatia. And, says Dieter Nentwich, head of the International Bureau in Jülich, this requires a strong science base.

The bureau, which is part of the German research ministry, currently supports around 15 scientific and technological collaborations with Croatia. These allow about 50 Croatian scientists at any one time work for a average of three months in a German research institute.

The bureau is now seeking a more direct influence on Croatia's research structure, and has offered to review the way its science is organized. According to Nentwich, the goal would be to make detailed recommendations on how research could be most efficiently funded, and on ways of encouraging the transfer of technology from laboratories to industry.

Croatia has not yet formally responded to Germany's offer. Croatian officials say they are likely to accept. But before doing so, they want to see the results of a similar exercise already being carried out in Slovenia, the wealthiest of ex-Yugoslavia's successor states's, which is expected by the end of the year. They also want to ensure strong representation of Croats on the proposed evaluation committee.

At present, uncertainties created by the most recent outbreak of fighting have led the bureau to hold back on its offer of help. Bureau officials say they are only prepared to act when the situation becomes "politically stable" — that is, when the war has ended.

But, even as Germany itself holds back, not everyone is standing still. The Rudjer Boskovic Institute, Croatia's largest research institute with 400 scientists working in its laboratories, is planning an internal reorganization. The German influence is clear. The institute's director, Krunoslav Pisk, says that its aim is "to be structured like a Max Planck Institute".

Robert Unterhuber

UNESCO falters in bid to draft human genome treaty

Paris. Next month, the general conference of the United Nations Educational Scientific and Cultural Organization (UNESCO) is due to consider proposals for an international agreement to control human genome research and its applications, a brainchild of its director-general, Federico Mayor.

But at a two-day meeting in Paris last week, the organization's International Bioethics Committee (IBC) failed to clarify either the legal status or the proposed content of such an agreement, fuelling sceptics' fears that, even if agreement is reached, it could be either ineffective or unworkable.

The committee is made up of 50 members drawn from both the developed and developing world. Almost a third are scientists, including three winners of the Nobel prize for physiology or medicine (Jean Dausset, Christian de Duve and Rita Levi-Montalcini) and one for chemistry (Sidney Altman).

Eight workshops were held in preparation for last week's meeting. But a planned debate of ethical issues never took place because several speakers ran over their allotted time. And work has not yet been completed by the organization's legal ex-

perts on the form of the accord.

Tom Forstenzer, a senior UNESCO official, defends the time spent on procedural matters as being necessary for people to get to know each other before beginning work. "They have to feel their way through different sets of laws and cultures."

But one member of the committee says that the meeting showed that the diversity of cultures represented at UNESCO makes the debate of ethical issues difficult, and could make consensus impossible. "We need to sort out issues at the national level first. Then countries with broadly similar cultures should see where to go from there."

The chairperson of the IBC is Noelle Lenoir, a prominent French bioethicist and author of the 1991 report that shaped the three bioethics bills submitted to the French Council of Ministers last year.

Lenoir strongly supports the organization's efforts to produce an international agreement on bioethics. UNESCO, she says, "combines the three aspects of the bioethical approach: science, education and culture." The need to take cultural diversity into account, she claims, will be the theme of the agreement. "It's a challenge, but I'm optimistic", she says. "We need to look at the sociocultural consequences of new technology and their effects on [North-South] equilibrium."

According to Lenoir, the IBC will probably draft guidelines rather than a full-blown treaty. The guidelines will be "concrete and pragmatic, and will not stigmatize research", she says. The committee plans to look at the responsibilities of genome researchers, "their right to accede to certain types of knowledge, and how they present results to the public".

The committee is likely to avoid one controversial topic, the ownership of genetic material. UNESCO has previously spoken out against the patenting of cDNA transcripts. But some committee members feel that it lacks sufficient expertise in patent law to reach a consensus.

UNESCO has been promoting the proposed treaty as the main aim of the IBC. But Lenoir says that this is a long-term objective. IBC's first task, she says, will be "to develop programmes to educate and inform teachers, researchers and decision makers, particularly in developing countries".

Clarifying its *raison d'être* may be IBC's biggest preoccupation before its next meeting. "Maintaining the committee simply as a forum — and to encourage participation by developing countries — would not be an unworthy task for UNESCO", says one committee member. "But it must claim to be doing more than that."

Declan Butler

India fails to lift research spending in industry

New Delhi. Liberalization policies introduced by the Indian government in the mid-1980s in an attempt to boost industrial spending on research and development have failed to achieve their objective, according to a study sponsored by the Department of Science and Technology.

Despite the government's encouragement, such spending has declined, while the dependence of Indian industry on technological imports has increased. As a result, says the study, conducted by the Centre for Technology Studies (CTS) in New Delhi, innovation continues to stagnate.

The process of liberalization was launched in 1985 and completed two years ago, when the government announced a package of measures designed to increase competition in the private sector.

These included the abolition of government controls over industrial enterprises and various steps to encourage foreign investment. Companies have been given freedom to expand, as well as to import technologies and to enter into foreign collaborations without government permission.

The government had hoped that these measures would force companies to increase their research and development spending in order to become more innovative — and thus more competitive.

But according to the CTS study, none of this has happened. Data collected from about 250 companies found that their spending on R&D fell from 1.1 per cent of turnover in 1980 to 0.8 per cent after the economic liberalization.

Faced with stiff foreign competition, many companies preferred to invest in advertising and marketing rather than in product improvement. The study also found that foreign equity in Indian companies had a negative effect on their own R&D efforts, as these companies tended to turn to their foreign collaborators for help.

Many companies cut back on R&D aimed at import substitution after import restrictions were lifted in 1991. No significant technology was acquired after liberalization as the companies lacked the money to make such purchases. Nor was there any evidence of companies becoming more innovative; indeed, the CTS study found that the number of companies taking out patents fell from 22 in 1985 to 13 in 1991.

Ghayur Alam, director of CTS, says that the findings of the study do not suggest that India should return to the restrictive policies of the pre-1985 era. "But they definitely [support] calls for a massive government financial support for R&D in industry, and for more industry-related research in government laboratories."

K.S. Jayaraman

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Ghilleen Prance, director of the Royal Botanic Gardens at Kew and formerly research director of the New York Botanical Garden, has been awarded the first annual prize of the International Cosmos Prize Committee. The prize, worth 40 million yen (\$375,000), is one of the richest in science, and comes from the profits made by the 1990 Expo in Osaka, Japan. Prance, an authority on the rain forests of South America, has said that he intends to use the money to continue his research following his retirement from Kew in 1997.

The tyranny of auditors

SIR — The US Secretary of Energy's decision to split the Superconducting Super Collider (SSC) project into separate construction and science components is ill-advised, no matter how politically expedient it may be (*Nature* **364**, 567; 1993). The building of a major facility is usually subject to compromise at some stage when decisions have to be made in response to some unexpected difficulty or cost overrun. Now it would seem that the accountants will have the final word when such decisions have to be made. Thus, when it is eventually completed, it is highly likely that the SSC will differ in some important respect from what the scientists would have chosen within the budgetary limits.

Nowadays, scientists' overriding concern seems to be not so much what they want to do but with what can be done with the available equipment that has not yet been done. In these circumstances, it should not be surprising that scientists will push for ever more expensive equipment as they strive for the competitiveness their auditors expect of them. I believe that our best hope of escaping from the spiralling costs imposed by the tyranny of the auditor lies with the politicians and other controllers of the purse strings who can see that the more constraints they impose on basic researchers the more predictable (and expensive) the results will be.

Signs are sometimes displayed in shops along the lines — "I have reached an agreement with the banks that they will not sell vegetables. In return I have agreed not to cash cheques." The Secretary of Energy's decision is probably irrevocable, but if the scientists promise not to audit perhaps she can extract a reciprocal promise from the accounts?

D. W. Braben

*Venture Research International Ltd,
13-14 Mount End,
Theydon Mount,
Epping, Essex CM16 7PS, UK*

Allain defended

SIR — We read with interest Declan Butler's News story (*Nature* **364**, 269; 1993) about the verdict of the French appeal court which has sent Jean-Pierre Allain to prison.

We have known Allain for three decades. From the early 1960s, when he prepared his medical thesis under the supervision of one of us, we recognized both his scientific and his human qualities. On his return from the United States in the mid-1970s, he was charged with the care of the largest concentration of French haemophiliacs at La Queue les-Yvelines.

In many instances, he took the haemophiliacs to his own home and he and his wife, Helen Lee, considered them as their own children. We are unable to imagine that Allain could be the Jekyll and Hyde character portrayed in the French press "*avide de sensationnel*". We cannot accept that what was a universal error, human in essence, could evolve into a deliberate crime.

Some other writers share the view that he didn't blow the whistle loudly enough to have attention paid to his warnings. In early 1985, those who held similar posts elsewhere shared Allain's uncertainties at that time.

In *Science* (**260**, 1262 & **261**, 422; 1993), it is stated that Allain ranked fourth in the world in citations per paper about research on AIDS. We sincerely hope that the French Supreme Court will decide to hear an appeal.

We share the views expressed by our British colleagues (*The Lancet* **342**, 232-233; 1993) and we look forward to Allain's prompt return to transfusion science and medicine in Cambridge.

Jacques P. Caen

*Institut des Vaisseaux et du Sang,
Hôpital Lariboisière,
Paris 75070, France*

Peter A. Castaldi

*University of Sydney,
Hospital Westmead,
Westmead, NSW 2145,
Australia*

SIR — Your leading article "Justice unevenly spread in Paris" (*Nature* **364**, 267; 1993) stresses the disadvantages of the sentence for the research community. But Garretta and Allain are primarily physicians who deliberately provided patients with infected material. They were brought to trial not as researchers or commercial managers, but as physicians who should have obeyed the time-honoured medical rule "*primum non nocere*" ("first of all, do no harm"). The lessons of this trial for the research community are therefore of minor importance compared with the hundreds and maybe the thousands of people contaminated by a deadly disease.

Badrig Mélékian

*6 rue Crevaux,
75116 Paris, France*

Man and beast

SIR — Williams *et al.* (*Nature* **364** 664; 1993) protest at the misuse of language in recent publications although they do not allege that the meaning is distorted by the misuse.

For my part, I take particular objection to Williams' misuse of 'animal', which derives from the Latin *anima*, meaning 'a soul', and should surely be reserved for those of God's creatures that have a soul

— humans. The King James's Bible, printed (1611) before this misuse gained currency, uses 'beast' for what Williams refers to as a 'nonhuman animal'.

In short, *tempora mutantur*, Mr Williams. He should realize that language changes, as do languages.

Oliver R. Dearlove

*Department of Anaesthesia,
Royal Manchester Children's Hospital,
Manchester M25 1HA, UK*

SIR — Williams *et al.* complain that words whose etymology implies human referents are now commonly used in nonhuman contexts. They even raise the possibility that such confusion may stem from a desire for political correctness masking the human/nonhuman distinction.

Two points are worth making. First, if political correctness is indeed an issue, it is important to remember that emphasizing such a distinction is as political as minimizing it. As to which is politically correct and which incorrect, I cannot say; but both are political.

Second, if ignorance is the issue, one might ask why it matters. Languages are dynamic, and meanings and usages change as the years go by. We no longer expect a sophisticated person to be 'impure' or 'adulterated', and the term 'sunrise' is commonly used by those who do not wish to imply that the Sun revolves around the Earth. If I read that there is an epidemic among farm animals, I know exactly what is meant; surely clarity and communicability are more important than etymological niceties, and we should avoid using two words when one will suffice to indicate our meaning?

R. Glynn Owens

*University of Wales, Bangor,
Health Studies Research Division,
Upper School,
St David's Hospital,
Bangor,
Gwynedd LL57 4SL, UK*

Not wind but water

SIR — I am sorry if I should appear to be attempting to sabotage your interesting essay on sand ripples (*Nature* **364**, 385; 1993) but although I have lived by the sea in various places most of my life, I do not actually know any beaches to which it would apply. Deserts, yes indeed, and I suppose also seaside sand dunes, but the typical ripples on the average holiday beach are produced under water and not by the action of wind. These are usually said to represent a standing wave pattern, and certainly it is easy enough to reproduce them in the bath if one has sandy enough feet.

David A. H. Taylor

*12 Avenue Road,
Scarborough, Yorkshire YO12 5JX, UK*

The memetic basis of religion

SIR — Both B. D. Josephson¹ and, earlier, A. Baidins² seek to find a way to study religion scientifically and they offer as a starting point the assumption that the ability to experience religious feelings may be encoded in the genes. Religion then is selected because “the central theme of religion is the attempt to maximize human goodness” and “because societies in which this potential is actualized . . . will tend to function more harmoniously and more efficiently”¹ or because “some humans are dimly aware of another dimension in this Universe . . . which helps them make more constructive decisions than those people lacking such a faculty”².

In my opinion, one should be very careful not to confuse different conceptual categories, such as atoms, genes, organisms, intelligence and thoughts.

In any case, I prefer to look at religion as an emergent characteristic, which can arise only after other levels have come to full development. Looking at it this way, religion has to do with the confrontation of our animal emotionality with our human superintelligence. I will try briefly to explain.

Animals have evolved as organisms that tend to reach what could be called ‘hormonal equilibrium’, which people call happiness. By performing certain tasks — feeding, fighting and sexual behaviour, which ultimately have the single purpose of multiplying the genes present — the animal is rewarded: it experiences joy, relief, satisfaction, or ‘happiness’. For the genes, this is at a different level in the conceptual hierarchy, an elegant, effective and universally applicable solution. Thus, regarding animals and humans as basically emotional beings is the first cornerstone in any hypothesis. The ability to enjoy and to fear underlies the evolutionary development of intelligence; animals and children learn through joy and fear, and experience emotional reward when making new discoveries.

The selective advantage of intelligence is easy to see: intelligence allows the storage of information from the past to tackle current situations more efficiently.

The superintelligence of human beings, however, is something else. Its capacity for making associations leads to unlimited fantasy and allows even speculation about the future. But that leads to a fundamental problem: animals fall asleep after a good meal or feel relieved after escaping a predator, but humans lose their ability fully to enjoy the present happiness because of anxiety over whether they will find food tomorrow or escape next time. And they do so on every possible occasion. Will our children grow happily? Will I pass my exam? Will I see her again?

What will happen when I die? For humans, instant animal anger and joy turn into endless fear and longing. Because mental events can strongly influence neuro-endocrinological functioning (and vice versa), humans become uncertain and are prone to depression.

Religious belief, in my opinion, is a human behaviour devised as a solution to this problem.

Logically one can do two things: try to solve the problems caused by our excess ability to make associations by more thinking and by constructing thoughts or theories that bring relief, or try to stop thinking (which is very difficult). The first solution gives rise to ideas that Richard Dawkins has called ‘memes’ — thought constructions that endow an individual with certainty about its own fate.

A memetic selection pressure can then be inferred: the most satisfactory memes will be selected from the thought pool and then distributed — horizontally — into different brains. Meme selection, meme classification and memetic evolution then become legitimate objects of study. Moreover, memes have a high adaptation capacity as is shown by the changes that Western religious doctrines have gone through to adapt to the findings of science.

So religion could be considered as a meme. Gods, divine powers, or holy trees are believed to let us influence the future. Indeed religion helps individuals to function more efficiently, as Baidins suggested². But that is not because some individuals have some genetic ability for religion, but because religion is a meme. Moreover, unlike Josephson and Baidins, I do not consider altruism as the essence of religion, but its memetic property as defined above. Initially people tried to influence gods and powers — and thus the future and their own fate — with offerings; only later were things such as promoting altruistic behaviour and rules for social organization grafted onto religion. Even then the essence of religion remains the influencing of the future: the major reason for behaving according to the rules is that one can propitiate the deity by doing so.

Another interesting meme, which could be regarded as co-evolving with religion, is anthropocentrism. That is a meme because it provides an individual with self-confirmation by the assurance that he/she belongs to a superb (divine) species.

To be fair, one should then ask whether science is a meme. Astronomy certainly was from the very beginning of cultural

history. Astronomy allows us to foresee events such as eclipses or seasonal weather changes. That may be why astronomy is the only branch of science developed to a high degree early in cultural history, and independently in different cultures. It brought power to those able to make the necessary calculations and was evidently considered sufficiently worthwhile to justify the building of enormous — scientific — measuring constructions. (Observe that priests were also astronomers in many cultures.)

For scientific reasoning itself to be accepted as a way of enhancing certainty had to wait for Newton. Then people suddenly realized that our own ‘reason’ could unravel divine rules. People started to ‘believe’ in science, which had been considered for many centuries as merely suited to resolve practical problems (such as how to build cathedrals for the purpose of religion, the major meme in the pre-scientific world). Thus the fact that science became a meme might be the basic reason for its sudden explosion.

The nonmemetic way to tackle the problem is to stop thinking. No one wants to eliminate the other root of the problem (feelings), because emotion developed earlier and is more essential to our functioning. Elimination of mental activity allows us to reach directly the state of animal happiness. Again, two major approaches can be distinguished.

Buddhism (generally not considered a religion, *pace* Josephson¹), is one technique. Meditation can produce a happy feeling of unawareness, the not-knowing state of animals.

Materialistic nonmemetic solutions become possible in societies of plenty. People try to stop thinking by emphasizing the essential (animal) needs: they try to stimulate their pleasure centres directly with all kinds of chemicals or they try to experience joy through evolutionary channels by exaggerated feeding, sexual or self-confirmation behaviour.

All of this is only a hypothesis, but one that allows the interdisciplinary scientific study of human behaviour, including religion. It might give us an opportunity to approach very different kinds of human behaviour (from killing for ideas to driving cars that are too big) with a single key: we are doing all of this because evolution made us naturally unhappy organisms, struggling with the emotional consequences of our excess of associative capacity.

Mario Vaneechoutte

Department of Clinical Chemistry,
Microbiology & Immunology,
Blok A, University Hospital,
B 9000 Ghent,
Belgium

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only.

1. Josephson, B. D. *Nature* **362**, 583 (1993).
2. Baidins, A. *Nature* **346**, 693 (1990).

Bringing the extinct dodo back to life

The dodo may long since have been extinct, but one zoologist has reconstructed the creature in sufficient detail for there to be evolutionary puzzles to be solved.

THAT the dodo deserves a better press is hardly anywhere disputed. While still extant, this flightless bird was both mocked and tormented by the sailors who found it marooned on Mauritius, in the Indian Ocean. Illustrations of an evidently ungainly animal were then taken as figures of fun by a cruel seventeenth century. A few specimens were captured on behalf of zoos, but did not long survive. And then the species (*Raphus cucullatus*) became extinct (around the year of Charles II's restoration to the British throne), except as a figure of speech (as in "As dead as a...") and in storybooks.

To set the record straight, one Bradley C. Livezy from the University of Kansas has now done his best to put the dodo on the sober basis any species deserves to be with a study labelled "An ecomorphological review of the dodo (*R. cucullatus*) and the solitaire (*Pezophaps solitaria*), flightless Columbiformes of the Mascarene Islands" (*J. Zool., Lond.* **230**, 247–292; 1993). The enterprise has been a mammoth undertaking, involving the use of data from more than 400 skeletal relics of the dodo, mostly at the Mauritius Institute. But that is only part of the puzzle of the dodo, a magnificently overweight pigeon.

Several questions arise. What, for a bird, are the benefits of not being able to fly? How did the dodo and the solitaire, distinct though they are, arise apparently independently on two islands that must be counted close neighbours by the yardstick of the Indian Ocean, but which are nevertheless nearly 600 km apart and thus too far for a flightless bird to make the passage? And why was the dodo such an ungainly creature?

In passing, Livezy inevitably provides his readers with a compelling comparison of the literary styles of the old and modern zoologists. Here, for example, is the description of the dodo due to H. E. Strickland and A. G. Melville in 1848:

These birds were of large size and grotesque proportions, the wings too short and feeble for flight, the plumage loose and decomposed, and the general aspect suggestive of gigantic immaturity.... So rapid and complete was their extinction that the vague descriptions given of them by early navigators were long regarded as fabulous or exaggerated, and these birds, almost contemporary of our great-grandfathers, became associated in the minds of many persons with the Griffin and Phoenix of mythological antiquity. The aim of the present work is to vindicate the honesty of the rude travellers of the 17th century....

Livezy first compares the two flightless birds with each other and with other

Columbidae. Of the two, the solitaire was the heavier bird — the average weight of males is estimated at between 20.9 and 27.8 kg, compared with 15.9 and 21.2 kg for the dodo. One striking feature of both species is the exaggerated sexual dimorphism; females were only two-thirds as heavy as males, and had much shorter bills.

It is clear from the comparisons of dodo and solitaire with other species that both are at the end of a spectrum. Altogether, representatives of 38 genera of Columbidae are used to show, for example, that while the wing-loading (in flight) of the other birds increased with overall mass, that based on the assumption that either the dodo or the solitaire could lift itself off the ground by flapping its undersized wings would be entirely off the regression line that accommodates the others. Dodos and solitaires were, perforce, perpetually grounded. They were too heavy to fly.

Interestingly, though, it seems that the adaptation of the bones of the legs and the pelvic girdle to a life on land is not nearly as great as that found in other flightless birds. To be sure, the femur is strong enough to support the whole weight of the bird, but it appears not to have been shortened in adaptation to the extent found elsewhere. The most marked differences are in the upper limbs, which are both lighter and shorter than in related birds, which is consonant with the absence of flight.

Livezy guesses at the lifestyle of the dodo on general physiological grounds. The physiology is simple allometry — the greater an animal's mass, the less will be its metabolic rate, and so on. But in the tropics, the dodo would have had to depend on seasonal fruits and other herbage accessible from the ground. So, the argument goes, each bird would have had to grow fat when food was plentiful to survive the privations of the thin season. That accords with accounts from the seventeenth century of the cycles of fattening and slimming in the dodo.

There are also contemporary accounts of the territorial behaviour of the bird. (Allometry has it that the bigger the bird, the bigger its territory and the more fiercely it is defended.) One early visitor to Mauritius described how dodos would defend their nests (made of a pyramid of palm leaves 18 inches high) to a range of 200 yards or so. Reproduction involved the laying of a single egg in the nest, which was incubated by each parent in turn. The young were slow to mature. The lifespan may have been as much as 30 years.

So how did two such similar species arrive at two islands in the same corner of a huge ocean? Livezy rejects the obvious notion that they belong to the same branch of a cladogram for the pigeon family that has not yet been constructed. It is not simply that there is no reasonable way of supposing that a pair of flightless birds would have made the 600 km crossing between the two islands, but also that the anatomical differences between them are no more impressive than their similarities with the Columbidae. And why not, anyway, allow that, if some small ancestral population of a species of the pigeon family can colonize an island and then abandon the capacity to fly for the evolutionary benefits of flightlessness, it might have happened a second time, or several times?

That is the fashionable way of thinking, but is not for that reason incorrect. In an island free from predators (as pristine Mauritius would have been), members of a primitive flighted population may have been able to survive even the disadvantage of being less strong flyers than their fellows, and may over the generations have followed the path of adaptation leading to the dodo in one case and the solitaire in the other. It may be too late to search the other islands of the Indian Ocean for surviving analogues of the dodo, but flightless birds are by no means uncommon — there are flightless rails in the Western Indian Ocean, for example.

Livezy's more intriguing evolutionary speculation picks up the phrase "gigantic immaturity" used by Strickland and Melville in their unavoidably second-hand account of the dodo. What did they mean by that? The untidy layout of a dodo's feathers, and its outer covering of down, for one thing. Elsewhere there is a description of a dodo as a "young duck or gosling enlarged to the dimensions of a swan", suggesting juvenility.

So was the dodo a consequence of paedomorphosis, the supposed phenomenon in which the development schedule of an organism is reorganized in such a way that it becomes sexually mature, and further development is halted, when some parts of its body have reached only an intermediate stage? That is all the more likely to happen when, from one generation to the next, the size of the individual members of a species is increasing. Livezy says of this as of many other questions about the dodo that perhaps we shall never know, but that may be too despondent; it is remarkable how much he has wrung from a few hundred bones.

John Maddox

No shortcuts in new maps

Peter Kareiva

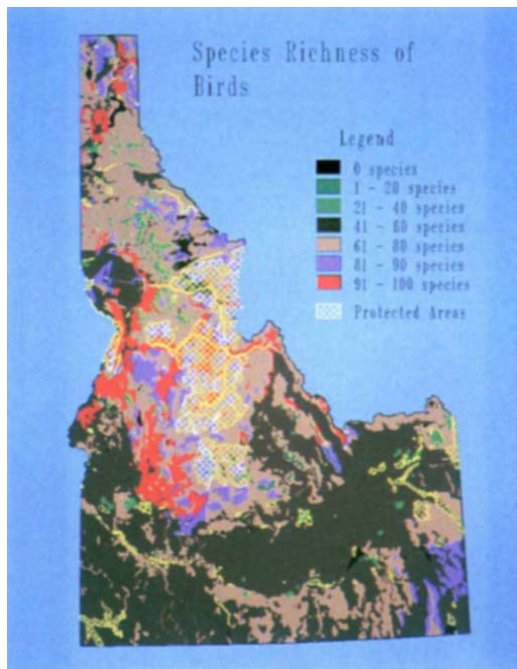
MORE than a million described and several million undescribed species inhabit our planet, so it should be obvious that we cannot fight to maintain biodiversity by taking up one species at a time. The futility of such an approach has led ecologists to begin exploring alternative strategies. One option is using distributional data on large numbers of species to identify areas that would not only provide refuge for endangered species, but would also be a haven for the majority of species not yet at risk. The first results from such an examination on a national scale are described on page 335 of this issue¹ by Prendergast and colleagues. Their report suggests that if data from Britain are any indication of worldwide patterns, there will be no easy shortcuts to biodiversity preservation.

All plans for the protection of biodiversity hinge on real estate — specifically designating large tracts of land to be protected from activities that cause species loss. As there are limits to how much land can be set aside in reserves, selecting the 'best pieces' is crucial to effective conservation. Unfortunately, these choices are typically made without many data and without adequate resources or time to conduct careful inventories. Conservationists have responded to the quandary by emphasizing two quick-and-dirty criteria for site selection: concentrations of species richness and the presence of rare (or endangered) species. The process could proceed remarkably well if concentrations of biodiversity for different taxa coincided, and even more so if such concentrations also overlapped regions with large numbers of rare species.

Prendergast and co-workers tapped into data describing the presence or absence of birds, butterflies, dragonflies, liverworts and aquatic plants in 10-km by 10-km squares throughout Britain. These data (held by the British Records Centre and the British Trust for Ornithology) provide species-by-species distributional records that are unrivalled by anything elsewhere in the world. Two elegantly simple questions were asked of the data. By protecting centres of diversity (hotspots) for one taxon, are we likely also to protect rich assemblages of others? And by protecting centres of diversity how likely are we at the same time to preserve parcels of land harbouring rare species?

The answers are discouraging. Areas species-rich for one group were generally unlikely to coincide with the species-rich

areas of other taxa; for example, bird and butterfly presence-or-absence data from over 2,500 of the 10-km by 10-km quadrats revealed that less than 12 per cent of the bird and butterfly hotspots coincided. In addition, many rare species cannot be found in any diversity hotspots, which means that land acquisition aimed exclusively at maximizing the number of species



A prototype gap analysis map in which the dominant vegetation of the state of Idaho was mapped at a scale of 1:250,000 using a minimum mapping unit of 200 hectares. Distributions for each bird species were then mapped using the vegetation types as map units. Bird species richness for each vegetation map unit is displayed in relation to conservation lands. The data came from satellite images, government agencies, herbaria and museum records, expert field biologists, air photographs and the published literature. Tests revealed that the map accuracy for vegetation type and species distributions were above 85 per cent. (Map courtesy of J. Michael Scott, director of the National Gap Analysis Project).

on protected lands is likely to miss uncommon species — the very ones most likely to be threatened by habitat destruction. Prendergast and colleagues cautiously point out that Britain is a unique island with a long history of habitat modification; it could be that the patterns evident in Britain reflect this extensive human influence and thus do not apply to wilder areas such as the tropics or even North America. Nonetheless, at this point we have to conclude that there is no evidence that simple rules of thumb can be used to guide strategies of land acquisition with the central goal of biodiversity protection.

This leaves open the question of

whether more detailed algorithms could yield better results when applied to maps of species occurrences². The most extensive mapping project in the service of biodiversity is being conducted in the United States. Here, the US Fish and Wildlife Service has initiated what it calls a Gap Analysis Program (GAP), in which dominant existing vegetation is mapped on a state-by-state basis and overlaid with data on actual species occurrences or predicted occurrences generated by habitat association models. GAP is underway in 28 states, and relies heavily on geographic information systems (GIS) to both codify and manipulate the data.

The project has an ambitious aim — to identify gaps in the overall mix of present-day conservation lands in the United States, so that limited land acquisition might provide protection for species and vegetation types currently left exposed³. At the heart of gap analysis is GIS computer technology, with all its promises and pitfalls. Because huge datasets are readily manipulated once entered into a GIS system, gap analysis can ask by exhaustive computer search what pieces of land would best serve biodiversity protection; for the British data this would correspond to asking which ten or twenty or thirty 10-km by 10-km quadrats out of the more than 2,000 quadrats in the database would best protect the most species, and how that land is currently managed. Such an approach contrasts sharply with posing a specific biological hypothesis about centres of diversity in the manner of Prendergast and colleagues.

The GAP is also distinguished by using as its starting point the distribution of vegetation types (often derived from satellite imagery and other records), which are hypothesized to predict animal occurrences. Specifically, points of collection for each animal species from museum records or government agencies are plotted on vegetation maps and then translated into measures of affiliation with different combinations of vegetation type, landform and climatic regime³.

Such habitat models are a necessity where one is concerned with large landscapes only sparsely sprinkled with direct observations (as is the case in most places, certainly most outside Europe). But before gap analysis is accepted, a great deal will have to be done to estimate errors and biases that arise from translating satellite data into vegetation types, then translating vegetation types into predictions about animal distributions, and finally using specific mapping units to identify priorities for biodiversity protection.

One promise of the GAP is that its answers can be dramatic and transpar-

ently practical. For example, an analysis of Idaho (see figure) revealed that although 70 per cent of the state is owned by the federal government, most of that land is not managed for biodiversity protection and a modest change in management could cover most of the vertebrate diversity in the state (J. M. Scott, personal communication). Of course, this leaves unanswered the dilemma raised by the study of Prendergast *et al.* — that preservation of vertebrates may leave butterflies, dragonflies and so forth out in the cold, because centres of diversity for different taxa do not seem to coincide. It would be unfortunate if we allowed our emotional affinity for vertebrates to short-change biodiversity as a whole (remembering that less than 5 per cent of known species are vertebrates, whereas half are insects). Scientists associated with the GAP recognize this pitfall, and have recommended giving greater attention to taxa such as butterflies, although efforts are hindered by a weak database.

The promise of biodiversity maps and species inventories has even been acknowledged by prominent politicians in the United States. Indeed, Bruce Babbitt, head of the Department of Interior (which includes the US Fish and Wildlife Service, the National Park System and the Bureau of Land Management), has made a national biological survey one of his priorities; a report from the National Research Council, on issues relating to the scope of the enterprise, is due to be published shortly. Babbitt advertises the survey as a solution to the “environmental train wrecks we see scattered across the country...”⁴, with the hope that such a survey will allow identification of critical resources most in need of protection before those resources are unwittingly squandered. Nearly 30 per cent of the land in the United States is federally owned, so it is clear that Babbitt’s initiative could reap enormous benefits without any intrusion on the public sector.

With both politicians and scientists turning towards species maps for illumination, there is no question that the study by Prendergast *et al.* is only the first of many important contributions. But inventories and maps alone will not solve any problems, or even point to the solutions — more work needs to be done testing hypotheses about patterns of diversity and exploring computer-based algorithms aimed at maximizing diversity. □

Peter Kareiva is in the Department of Zoology, University of Washington, Seattle, Washington 98195, USA.

An elementary history lesson

M. G. Edmunds

MAN’S fate may not be written in the stars, but the history of the synthesis of the elements is certainly decipherable from their chemical compositions. A considerable advance for historians of our Galaxy comes with the appearance of a long-awaited study of the detailed elemental composition of 189 dwarf stars in the solar neighbourhood¹. The work represents some 10 years of coordinated observation and analysis by a team of six astronomers, making a delightful change from the “we have observed one object. . .” school of hit-and-run astronomy.

The study is based on F- and G-type dwarf stars, types that are not very different from our own Sun. Indeed, the analysis shows the Sun to have a ‘typical’ elemental composition. The sample was selected from previous photometric surveys to give stars whose ages could be estimated, spanning more or less the whole history of the galactic disk, and a range of a factor of 20 in heavy element abundance. It is the increase in precision of measured abundances that makes this study so telling. Once, it would have been inadvisable to trust a stellar abundance analysis to better than a factor of two — and those were the good ones — but with higher-sensitivity detectors on modern spectrographs, and more sophisticated models of stellar atmospheres, Edvardsson and collaborators¹ can give abundances and abundance ratios accurate to about 12–25 per cent. Other analyses have approached such accuracies, but never in such a large and carefully chosen sample. With the errors beaten down, real trends, suspected from previous studies, become firmly established.

The abundance of iron is traditionally used as a measure of the overall heavy-element abundance (‘metallicity’) of stars, largely because it shows many spectral lines in the F-, G- and K-type stars, which for technical reasons are the most easily analysed. To a first approximation, strongly qualified below, the old disk stars in the solar neighbourhood show low metallicity and the young stars show high metallicity. The ratios of other elements to iron as a function of overall metallicity can then be a useful probe of past nucleosynthesis processes².

The so-called α -elements, magnesium, silicon, calcium and titanium, are built up in stars and their supernovae by, in effect, the successive addition of helium nuclei. Like oxygen, they show a higher abundance relative to iron in low-metallicity (old) stars than in high-metallicity (young) stars. This has been taken to indicate that when element synthesis started with the release of newly synthesized α -elements

and oxygen from (type II) supernova explosion of short-lived massive stars, there was then a delay until the type Ia supernovae from longer-lived, lower-mass binary systems contributed the iron-group elements. Edvardsson and co-workers have identified slight variations in the behaviour of the α -elements — magnesium, for example, shows more variation than calcium — which will be seized on as valuable tests of both galactic evolution and nucleosynthesis theory. An encouraging aspect is the apparent decrease in α -element excess relative to iron with increasing galactic radius. This can be interpreted as implying that the inner regions of the Galaxy formed more quickly than the outer parts, and agrees with preliminary observations³ of high magnesium-to-iron ratios in stars in the galactic bulge. It may also confirm suspicions⁴ that magnesium has a high relative abundance near the centres of many galaxies whose formation might be expected to have been rapid.

Edvardsson and co-workers’ sample shows that the metallicity of the galactic disk has built up slowly during its history, following something between a square-root and a linear increase with time. The metallicity in the gas of the disk decreases with increasing radius⁵, and a comparable radial decrease is seen in the progenitor material of these stars, once their birthplaces have been estimated from their present-day motions.

A striking result is the scatter in the metallicities of stars formed at similar times at the same distance from the centre of the disk — it is almost comparable with the overall increase in metallicity with time. We could once have blamed analysis errors for such a spread, but now it seems that it is real. However, the heavy element ratios show less scatter, implying that all the different nucleosynthesis sites contributing different elements to the initial abundances in the stars mix their products together well. This seems to exclude the possibility that inhomogeneities build up because stars form in regions strongly polluted by the products of just one or two supernovae. Edvardsson *et al.* suggest that the metallicity spread may result from inhomogeneous accretion of metal-poor gas from outside the disk, a popular ingredient in galactic chemical evolution models. But it is difficult to see why the processes that mix the nucleosynthesis products into the disk would not thoroughly mix any accreted gas as well.

The most noteworthy of the other element ratio trends are those, like Sherlock Holmes’s dog that did not bark in the night, that do not vary. Sodium was

1. Prendergast, J. R., Quinn, R. M., Lawton, J. H., Eversham, B. C. & Gibbons, D. W. *Nature* **365**, 335–337 (1993).
2. Pressey, R. L. *et al.* *Trends Ecol. Evol.* **8**, 124–128 (1993).
3. Scott, J. M. *et al.* *Wildl. Monogr.* **123**, 1–41 (1993).
4. Babbitt, B. *The National Biological Survey* (US Department of Interior, 8 June 1993).

expected (on nucleosynthesis grounds) to vary strongly with metallicity, but stubbornly remains in almost constant ratio to iron. Barium and other elements formed by slow-timescale neutron capture processes in giant stars do not show the expected rapid increase in abundance with increasing metallicity. The reason may be⁶ that increasing metallicity effectively poisons and reduces the bombarding neutron flux, in rough compensation for the increasing availability of heavy target nuclei.

Modern stellar abundance analysis has real potential. With efficient high-dispersion spectrographs on our largest telescopes, and the future introduction of even larger telescopes, this sort of investigation could be extended to regions of our Galaxy (and perhaps eventually of

other galaxies) far beyond our local solar neighbourhood. In the mean time, the chemical evolution theorists will be busy modifying their models to account for the feast of observational trends provided by Edvardsson and collaborators. □

M. G. Edmunds is in the Department of Physics and Astronomy, University of Wales College of Cardiff, Cardiff CF2 3YB, UK.

1. Edvardsson, B. *et al. Astr. Astrophys.* **275**, 101–152 (1993).
2. Wheeler, J. C., Sneden, C. & Truran, J. W. *A. Rev. Astr. Astrophys.* **27**, 279–349 (1989).
3. McWilliam, A. & Rich, R. M. *Astrophys. J. Suppl.* (in the press).
4. Worthey, G., Faber, S. M. & Jesus Gonzales, J. *Astrophys. J.* **398**, 69–73 (1992).
5. Shaver, P. A. *et al. Mon. Not. R. astr. Soc.* **204**, 53–112 (1983).
6. Clayton, D. D. *Mon. Not. R. astr. Soc.* **234**, 1–36 (1988).

RNA PROCESSING

Splicing in stereo

Angus I. Lamond

AN intriguing question for those interested in splicing is whether the introns in precursor transcripts of messenger RNA (pre-mRNAs), which are removed by spliceosomes, share an evolutionary relationship with self-splicing RNAs found in mitochondria, chloroplasts and bacteria. Experiments reported by Moore and Sharp (page 364 of this issue¹) provide insights into the stereochemistry of the pre-mRNA splicing mechanism, and may lead to a more rigorous comparison of pre-mRNA and self-splicing introns.

Spliceosomes catalyse intron removal in the nucleus, using a reaction mechanism that involves two distinct transesterification steps. The first of these exchanges the 3'–5' phosphodiester bond at the 5' exon–intron junction for a 2'–5' phosphodiester bond formed between the 5' terminus of the intron and a ribose 2' hydroxyl group at a specific position within the intron. The second involves a direct swap of 3'–5' phosphodiester bonds, resulting in ligation of the exons and displacement of the intron. So far as we know, all pre-mRNA introns are spliced by a similar mechanism. The structure of the spliceosome is also very similar throughout the eukaryotes — in both yeast and mammals, spliceosomes are composed of three RNA/protein subunits, the U1 and U2 small nuclear ribonucleoproteins (snRNPs) and the [U4/U6.U5] tri-snRNP, together with another group of non-snRNP protein-splicing factors. By means of some particularly elegant biochemistry, Moore and Sharp have now been able to probe the active site of the mammalian spliceosome and to define the stereochemical course of each of the two transesterification reactions.

Making use of the different properties of diastereomers to probe chemical reaction mechanisms is a well-established procedure. It is not surprising, therefore, that the stereochemistry of the transesterification reactions in the spliceosome can be analysed by comparing how splicing of pre-mRNA substrates is affected by introducing a single R_P or S_P phosphorothioate linkage at either the 5' or 3' intron–exon junction. The tricky part is finding a way to prepare the requisite modified substrates. Although modified groups can be introduced readily at specific sites in short RNA molecules using direct chemical synthesis, the comparatively large RNA molecules needed for splicing substrates (usually several hundred nucleotides or more) cannot be prepared in a single synthesis by current methods. Conversely, larger RNAs can be made easily by *in vitro* transcription of DNA templates. But in this case modified or labelled groups become incorporated randomly throughout the molecule and cannot usually be targeted to unique positions.

Last year, Moore and Sharp came up with a clever solution to these difficulties². Their method involves first preparing two or more shorter RNA molecules, either by chemical synthesis or by *in vitro* transcription. The desired modifications are introduced into these shorter RNAs and the separate pieces then joined together with T4 DNA ligase, using a DNA bridge to hold the RNAs together (see figure).

Using this method, Moore and Sharp constructed pre-mRNAs with a single phosphorothioate linkage, corresponding to either the R_P or the S_P diastereomer¹. At both the 5' and 3' splice sites, splicing was completely blocked by the presence of

the R_P, but unaffected by that of the S_P diastereomer. This differs from group I self-splicing introns, where only R_P is accepted for the first step and S_P for the second^{3–5}. The differential effect of the phosphorothioate diastereomers on pre-mRNA splicing may reflect an essential interaction between one or more magnesium ions and one of the non-bridging oxygens in the scissile phosphate, because sulphur is a much poorer ligand than oxygen for magnesium. This is consistent with a model for splicing (and other RNA-catalysed phosphotransfer reactions) involving a catalytic site with two, closely spaced, divalent metal ions, analogous to that found in certain protein phosphoryl transfer enzymes, such as alkaline phosphatase⁶.

Because the S_P phosphorothioate diastereomer did not block either step of pre-mRNA splicing, it could also be used to examine the configuration of the scissile phosphate after transesterification had occurred¹. The S_P and R_P configurations could be distinguished by cleaving the RNAs with snake venom phosphodiesterase, which splits both 3'–5' and 2'–5' phosphodiester bonds but is much more active on the R_P than the S_P phosphorothioate diastereomer. Analysis of appropriately modified substrates, containing exclusively S_P diastereomers, showed that after each step of splicing the single phosphorothioate linkage became sensitive to cleavage by phosphodiesterase. This clearly indicated that the S_P had been inverted to the R_P configuration at each splice site, as a consequence of the transesterification reaction.

Although more complex schemes cannot be excluded, by far the simplest mechanistic interpretation of these data is that each reaction occurs as a single, in-line S_N2 nucleophilic displacement, not involving the formation of any covalent intermediates. As the authors point out, the data also indicate that the two steps of pre-mRNA splicing cannot be catalysed as forward and reverse reactions by the same active site, which contrasts with the situation for group I self-splicing introns. So there must be different active sites in the spliceosome for the two transesterification reactions (although it remains to be seen whether the splicing intermediates are translocated to a new active site in the spliceosome, or, as is

1. Moore, M. J. & Sharp, P. A. *Nature* **365**, 364–368 (1993).
2. Moore, M. J. & Sharp, P. A. *Science* **256**, 992–997 (1992).
3. McSwiggen, J. A. & Cech, T. R. *Science* **244**, 679–683 (1989).
4. Rajagopal, J., Doudna, J. A. & Szostak, J. *Science* **244**, 692–694 (1989).
5. Suh, E.-R. & Waring, R. B. *Nucl. Acids Res.* **20**, 6303–6309 (1992).
6. Steitz, T. A. & Steitz, J. A. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6498–6502 (1993).
7. Green, M. R. *A. Rev. Cell Biol.* **7**, 559–599 (1991).
8. Sharp, P. A. *Cell* **42**, 397–400 (1985).
9. Cech, T. *Cell* **44**, 207–210 (1986).
10. Weiner, A. *Cell* **72**, 161–164 (1993).

perhaps more likely, whether a rearrangement or modification of the first active site takes place after the first step).

This two-site model for the spliceosome fits well with earlier genetic and biochemical data indicating that independent recognition of the splice-site sequences takes place at each step^{1,7}. For example, a specific block of the second splicing step (resulting in the accumulation of splicing intermediates) occurs with certain point mutations in either the branch point or splice sites of the intron, or in U2 or U6 snRNAs.

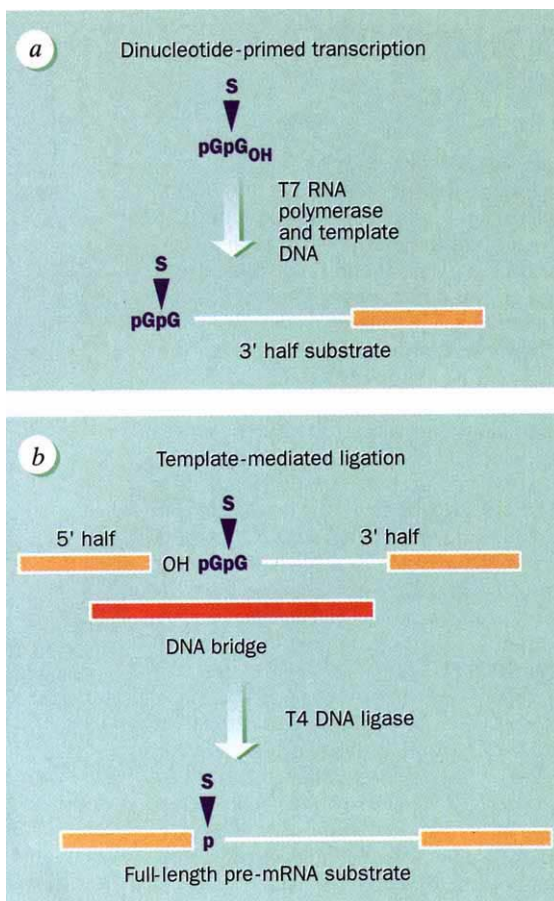
A similar phenotype can also be observed in yeast strains that have mutations in genes encoding splicing factors, or in mammalian splicing extracts treated with the protein phosphatase inhibitor okadaic acid. It is possible that in at least some of these cases the phenotype may arise because of a failure to switch the RNA intermediates in the spliceosome to the active site required for catalysis of the second step.

It will be of great interest now to probe the stereochemistry of the mechanism used by group II self-splicing introns. Unlike pre-mRNA introns, group II introns self-excite and do not require *trans*-acting factors. But they do use a two-step transesterification mechanism that generates identical RNA intermediates and products to those formed in the spliceosome, and it is tempting to speculate that there is an evolutionary relationship between pre-mRNA introns and contemporary group II self-splicing introns — that is, both may be descended from a common ancestral intron, but have diverged such that the intramolecular RNA base-pairing interactions that structure self-splicing introns are provided in *trans* for pre-mRNA introns by the spliceosomal snRNAs^{8,9}. Evidence pointing to striking similarities between interaction domains in self-splicing introns and base-pairing interactions detected in the spliceosome is consistent with, but does not prove, this model — as discussed by Weiner¹⁰, these similarities may reflect a process of chemical determinism rather than reflecting an evolutionary relationship.

relationship.

If future analyses show that pre-mRNA and group II self-splicing introns have identical stereochemistry, the issue remains open. But if group II introns accept the R_p phosphorothioate diastereomer, at either splice site, we will know that they are not in fact related to pre-mRNA introns after all. Given the recent history of surprises in the RNA field, I am eagerly awaiting the answer and making no predictions. □

Angus I. Lamond is in the European Molecular Biology Laboratory, Meyerhofstrasse 1, Postfach 102209, D-6900 Heidelberg, Germany.



The template-mediated ligation method², for introducing modifications at preselected sites in RNA, is illustrated here for insertion of a single phosphorothioate at the 5' splice site in a pre-mRNA¹. *a*, The 3' half-molecule is made by dinucleotide-primed, *in vitro* transcription of a DNA template containing the desired sequence adjacent to a phage polymerase promoter. The dinucleotide containing a single internucleotide phosphorothioate was made by chemical methods and the 5' phosphate was added enzymatically. The R_p and S_p phosphorothioate diastereomers can be separated by chromatography of the dinucleotides and hence used to prepare separate 3' half-molecules with either configuration. *b*, The corresponding 5' half-molecule (also prepared by *in vitro* transcription) is joined to the 3' half-molecule using T4 DNA ligase, with the two RNAs precisely aligned for ligation on a single-stranded DNA bridge spanning both sequences. The product is a full-length pre-mRNA substrate containing a single S_p or R_p phosphorothioate linkage at the 5' intron-exon junction.

RÉSUMÉ

Beaver boom

THE spectre of population bottlenecks looms large for conservationists; the worry is that after a population crash, loss of genetic variation and inbreeding depression may compromise the species' recovery. H. Ellegren *et al.* have come up with heartening news on that score (*Proc. natn. Acad. Sci. U.S.A.* **90**, 8150–8153; 1993). By the mid-nineteenth century the beaver had disappeared from Sweden and its numbers stood at maybe a few tens of individuals in Norway. But introduction of animals from Norway into Sweden in the 1920s and 1930s paid off, for Swedish beavers now number around 100,000. Ellegren and colleagues find that this population explosion occurred despite unusually low genetic variability, including that in the major histocompatibility complex. The results show that limited MHC polymorphism does not necessarily compromise recovery of an endangered species (at least over short ecological timescales).

Greater crater

Of the suggested causes of the mass extinctions at the end of the Cretaceous and start of the Tertiary era 65 million years ago, a giant meteorite impact is probably the front runner. The buried geological structure at Chicxulub, Mexico, is thought to be the remnant of the impact crater.

V. L. Sharpton and collaborators, writing in last week's *Science* (**261**, 1564–1566; 1993), now say that the diameter of the structure may be far greater than the earlier estimate of 200 km. Beyond the three concentric circles detected in earlier gravity maps, they find broken arcs of a fourth. If these mark the outer rim of the original crater, it was huge: about 300 km across, making it the site of one of the largest known impacts on any planet in the inner Solar System in the past four billion years.

Turning yellow

NOT least among the trials of advancing decrepitude is the increasing opacity of the lens. In severe cataracts this results from loss of the short-range structural order, to which the lens owes its transparency. Now G. Suárez *et al.* (*J. biol. Chem.* **268**, 17716–17721; 1993) have examined normal lenses from people between 6 and 82 years old by X-ray scattering: it turns out that scattered intensity increases rather abruptly from the age of about 55. This parallels, and by implication is caused by, accretion of a fluorescent, urea-insoluble fraction in the lens protein. Such cross-linked aggregates are secondary products of nonenzymic glycation by the Maillard reaction (formation of Schiff bases from reducing sugars and protein amino groups), which generates characteristic fluorophores and yellowing. The melancholy message may be to lay off the chocolates (as well as everything else).

The information explosion

Tony Stace

How can we capture the instant at which a molecule undergoes a chemical reaction? It's a tall order: a molecule can take just 10^{-13} seconds to react, during which its geometry may change extensively. Femto-second transition-state spectroscopy, where extremely short light pulses are used to track changes in nuclear motion as molecules react, has been hugely successful¹. But spectroscopic methods are not always appropriate when rapid changes in molecular geometry occur and rigid stable molecules, held together with strong chemical bonds, become floppy reaction intermediates with weak binding energies. Now Vager and co-workers², whilst using a very different technique known as Coulomb explosion to capture stable modes of molecular behaviour, have as a bonus observed an intermolecular rearrangement in action.

The Coulomb explosion technique translates molecular dimensions into macroscopic images; whereas X-ray and electron diffraction do this by using the static amplifying properties of reciprocal space, Coulomb explosion uses real space and time³. A bunch of molecular ions of a

single mass, such as CH_4^+ or C_3^+ , is accelerated to more than 2 per cent of the speed of light (about 200 keV per a.m.u.). At this point the ions pass through a very thin (about 30-Å) metal target that strips away all the binding or valence electrons with minimum distortion to the geometry of the molecular ion. Emerging from the foil in close proximity, the resulting atomic ions (C^+ , H^+ and so on) suffer strong Coulomb repulsion and fly apart. To identify the charge-to-mass ratios of the atomic ions, they are deflected in a horizontal electric field downstream from the foil, finally striking a position- and time-sensitive detector a further 6 metres away (see Fig. 1a).

Figure 1b shows the detected pattern of fragments from the Coulomb explosion of about 10^4 CH_4^+ ions³. In order of increasing deflection, the positions of C^+ , C^{2+} , C^{3+} and C^{4+} can be recognized, together with a contribution from accompanying protons. The 10^4 explosions ('events') identified by the detector are analysed in terms of arrival time and position to give each molecular ion a unique coordinate in what is termed velocity space.

Consider the simple example of a homonuclear diatomic ion: if the ion explodes with its molecular axis parallel to the face of the detector, then both fragments will arrive at the same time but at different positions. If the axis is perpendicular to the detector then the fragments will strike the same spot but at different times. So each event produces fragments with a unique set of position and time coordinates which can be

equated to a velocity vector originating from the point of explosion (the foil). Numerical techniques are used to transform each event from velocity to coordinate- or R -space, which for CH_4^+ means nine structural degrees of freedom with respect to the centre of mass.

In CH_4^+ , the representation in R -space of all the H-C-H bond angles tells us both the symmetry of the ion and the reaction path. Neutral methane is tetrahedral, but once an electron is removed the resultant ion becomes distorted by the Jahn-Teller effect, which forces nonlinear molecules and ions in degenerate electronic states to adopt less symmetrical non-degenerate forms. CH_4^+ has C_{2v} symmetry⁴ — equivalent representations of the ion can be generated by reflections through two planes perpendicular to each other. The distorted CH_4^+ can adopt a series of

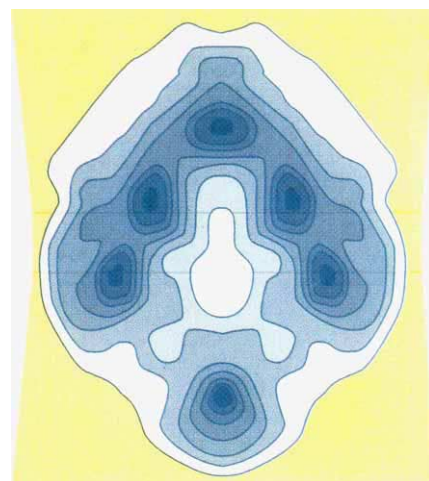


FIG. 2 Composite view of the image generated for CH_4^+ from the experimental Coulomb explosion data. The regions of high density correspond to ions with C_{2v} symmetry and these are connected through ridges created by ions that were undergoing pseudorotation (hydrogen atom permutation) at the time of the explosion. From ref. 2.

equivalent conformations which are connected by permutation of the hydrogen atoms. The change from one conformation to another can be achieved simply by rotation or by the hydrogen atoms tunnelling through a small energy barrier (about 80 meV high), a process often referred to as pseudorotation. What distinguishes rotation from pseudorotation, or reaction, is that reaction proceeds via a lower-symmetry structure, C_s , which has just a single plane of reflection.

What Vager *et al.*² have done is to identify, from their 10^4 recorded Coulomb explosion events, the presence of some structures with C_s symmetry. These configurations are part of the ensemble created as a result of ionization and then captured as an image at that instant in time when the hydrogen atoms are undergoing permutation. Figure 2 shows a composite representation of the geometry of all the captured images, derived from data like those in Fig. 1. Each of the regions of high density represents a structure with C_{2v} symmetry; but connecting these are ridges that can be identified as reaction paths as a consequence of pseudorotation. All the CH_4^+ structures detected are distorted; the coordinate system is such that an ion of tetrahedral symmetry would generate a dense region at the centre of the diagram³.

In part, the success of these experiments stems from the use of supersonic cooling to reduce the vibrational temperature of the ions. Vibrational motion can distort the equilibrium geometry of an ion and thus cloud the very subtle effects responsible for the features shown in Fig. 2.

Unlike other (spectroscopic) methods of studying chemical reactions¹, Vager

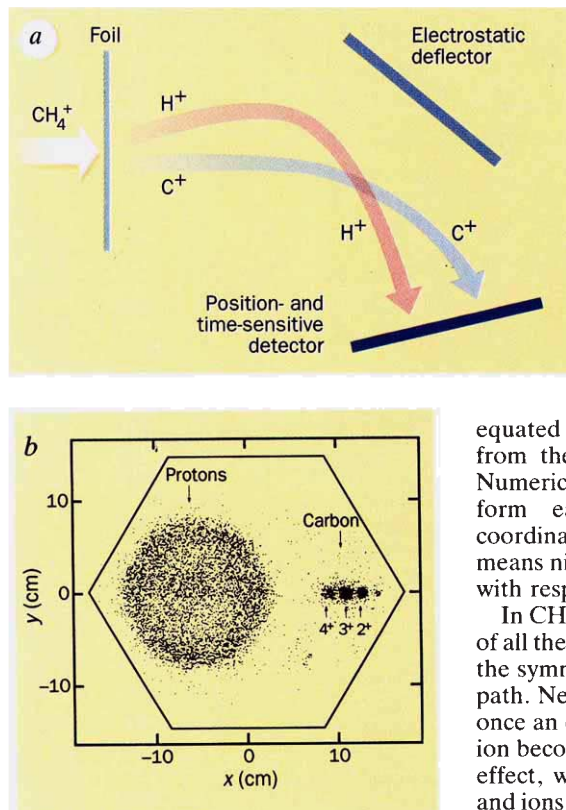


FIG. 1a, How the Coulomb explosion technique works. b, A two-dimensional density plot recorded at the detector of the images of fragment ions resulting from the Coulomb explosion of CH_4^+ ions.

and co-workers' results have no element of time resolution. The experiment works because the events of interest happen to be there at the instant when the ion is stripped of its electrons. At the same time, we should not underestimate the experimental effort needed to make such measurements.

To create images of reaction pathways or trajectories with femtosecond time resolution, Williamson and Zewail⁵ have recently proposed the use of ultra-short (femtosecond) electron pulses to generate diffraction patterns from chemical intermediates. As diffraction patterns are sensitive to internuclear separation and atomic number, they could be used to track changes in structure during a reaction and would eliminate the problem of having to interpret the spectroscopy of a rapidly changing chemical conformation.

Whichever technique proves to be the more versatile, the experimentalists must ultimately come up against the Uncertainty Principle; they may have an accurate picture of the position or time of a chemical event, but this is at the expense of knowing the momenta or energies of the atoms concerned. □

Tony Stace is at the School of Chemistry and Molecular Sciences, University of Sussex, Brighton BN1 9QJ, UK.

1. Khundkar, L. R. & Zewail, A. H. A. *Rev. phys. Chem.* **41**, 15–60 (1990).
2. Vager, Z., Graber, T., Kanter, E. P. & Zajfman, D. *Phys. Rev. Lett.* **70**, 3549–3552 (1993).
3. Vager, Z. in *The Structure of Small Molecules and Ions* (eds Naaman, R. & Vager, Z.) (Plenum, New York, 1987).
4. Frey, R. F. & Davidson, E. R. *J. chem. Phys.* **88**, 1775–1785 (1988).
5. Williamson, J. C. & Zewail, A. H. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5021–5025 (1991).

PLANT GENETICS

Pastoral synteny

Robert Shields

It is a sobering thought that the size of the human genome is the same as that of maize and lies midway between that of potato and pea. This illustrates the well-described but ill-understood C-value paradox, whereby the DNA content of many multicellular eukaryotes bears little relationship to their complexity. This is especially true of plants; the genome size of bread wheat is a hundred times that of *Arabidopsis*, and the tulip genome is larger still. Now, thanks to comparative genetic linkage mapping, it is becoming clearer how plant genomes got to where they are now.

The most extensive linkage maps in plants are based on restriction fragment length polymorphisms (RFLPs) detected using cloned DNA sequences as probes. Many of these probes are fragments of genes, so the RFLP map gives the relative

gene order in the chromosome. The early mappers worked in isolation on their favourite plant, using their own set of probes, and so maps were difficult to compare even between laboratories working on the same crop. But it is now emerging that quite diverse crops have related genome organization.

The first example was the demonstration, using probes isolated from tomato, that the maps of tomato and potato showed essentially the same gene order (synteny) on each of the 12 chromosomes that make the haploid set of both plants¹. It was therefore a bit of a disappointment that pepper (which is closely related to tomato, and has the same number of chromosomes but a three-times larger genome) showed little conservation of gene order².

Things looked rather brighter in the

cereals. Wheat, barley and rye are members of the family Graminae and are known on cytogenetic and other grounds to have chromosomes that are closely related³. They have the same basic chromosome number ($n=7$) and roughly the same DNA content per chromosome set (see box). So perhaps it was not too surprising that they revealed extensive synteny when mapped by RFLPs^{4–6}.

A bigger surprise came with the demonstration that large regions of the sorghum genome had the same gene order as maize — although both maize and sorghum come from the same botanical tribe, and have the same chromosome number ($n=10$), the maize genome is 3.5 times larger⁷. More extensive mapping reveals a number of inversions, relative translocations and possible loss of gene segments, but the general pattern is consistent with a common ancestor with $n=5$ which underwent polyploidization before the separation of the maize and sorghum lineages⁸.

Two new papers from a group led by Steven Tanksley have pushed the comparative genome analysis of the cereals a stage further^{9,10}. The first compares rice with maize. Although both are members of the Graminae, these species are in different tribes, have a different number of chromosomes and the maize genome is six times that of rice. But it turns out that the two genomes have extensive conservation of gene order over large segments (although relative translocations and duplications of a rice-like genome are seen in maize), and in some instances entire chromosome arms are nearly identical in gene order and content. Not satisfied with this, Tanksley's group extended the analysis to wheat, and the same pattern of widespread synteny was repeated. The most recent results, described by N. Kurata at a conference in Birmingham, UK, suggest that each wheat chromosome group is colinear with a chromosome of rice.

What all this reveals is that the Graminae have genomes that are related, probably by descent from an ancestral progeni-

Genome sizes of some of the Graminae

Species	2n chromosome number	Genome size (1C) pg	Mbp/1C	cM	kb/cM
Rice	24	0.45	430	2,300	191
Sorghum	20	0.8	800	1,530	191
Maize	20	2.6	2,500	2,200	1,136
Barley	14	5.5	4,900	1,403	3,471
Rye	14	7.9	7,600	>1,000	
Wheat	42	16.55	16,000	6,300	2,539
(2n = 6x)					

The size of the genomes are given in both picograms (pg) per chromosome set (1pg = 965 million base pairs (Mbp)), and in recombination units, where 1 centimorgan (cM) is a frequency of recombination of 1 per cent. 1C, unreplicated haploid genome.

1. Tanksley, S. D. *et al. Genetics* **132**, 1141–1160 (1992).
2. Tanksley, S. D., Bernatzky, R., Lipitan, N. L. & Price, J. P. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6419–6423 (1988).
3. Monte, J. V., McIntyre, C. L. & Gustafson, J. P. *Theor. appl. Genet.* **86**, 649–655 (1993).
4. Devos, K. & Gale, M. *Outlook on Agriculture* **22**, 93–99 (1993).
5. Devos, K. M. *et al. Theor. appl. Genet.* **85**, 673–680 (1993).
6. Devos, K. M., Millan, T. & Gale, M. D. *Theor. appl. Genet.* **85**, 784–792 (1993).
7. Hulbert, S. H., Todd, T. E., Axtell, J. D. & Bennetzen, J. L. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4251–4255 (1990).
8. Pereira, M. G. *et al. Genome* (in the press).
9. Ahn, S. & Tanksley, S. D. *Proc. natn. Acad. Sci. U.S.A.* **90**, 7980–7989 (1993).
10. Ahn, S., Anderson, J. A., Sorrells, M. E. & Tanksley, S. D. *Molec. gen. Genet.* (in the press).
11. Moore, G., Gale, M. D., Kurata, N. & Flavell, R. B. *BioTechnology* **11**, 584–589 (1993).
12. O'Brien, S. J. & Graves, J. A. M. *Cytogenet. Cell Genet.* **58**, 1124–1151 (1991).
13. O'Brien, S. J. *et al. Nature Genet.* **3**, 103–112 (1993).
14. Leitch, I. J. & Heslop-Harrison, J. S. *Genome* (in the press).

tor genome, and that relative gene order has often been conserved despite vast increases in genome size¹¹. Conservation of gene order should have been expected by those who follow the comparative gene mapping effort in mammals, where extensive colinearity is seen in genomes as diverse as those of humans, cows and even the platypus¹². However, the genome size variation in the Graminae is much larger than that in mammals (the entire rice genome could fit in a single wheat chromosome).

It seems that much of the DNA of the Graminae chromosome is made up of repeated sequences amplified to different degrees; indeed most of the size difference between the species is accounted for by

such sequences which have diverged faster than the coding regions. The mystery is how this amplification has been accommodated without destruction of the basic genome organization. It is likely that the entire genome does not undergo uniform expansion, but that the expansion may occur preferentially around the centromeres (see box).

The relationship between cereal genomes is of more than evolutionary interest. For instance loci governing similar plant characteristics can be seen in conserved regions of the genomes of rice, wheat and maize, suggesting that they may be encoded by different alleles of the same gene¹⁰. So the study of genetic variation in one species can tell us something about how phenotype is controlled in another. It may also be possible to

target genes for positional cloning in a species with a large genome, by genome walking in a species such as rice which has a small one.

How far comparative mapping can be taken is unclear. Members of the Brassicaceae are known to have similar genomes and work is underway to relate these genomes to that of *Arabidopsis* (another crucifer). If maps are to be compared in different crops and between different laboratories, the plant community will have to follow the lead of the mammalian gene mappers and agree on a set of anchor loci¹³. Only then will the full potential of comparative mapping be realized. □

Robert Shields is at Plant Breeding International, Maris Lane, Trumpington, Cambridge CB2 2LQ, UK.

PLANT genome maps are constructed by crossing plants that differ in the length of restriction fragments revealed by a DNA probe (an RFLP) and examining the inheritance of the RFLPs in the offspring. The order and distance between the RFLP markers that make up the genome map are determined by the frequency of recombination of the markers inherited from the parents. Physically close markers are assumed to recombine less frequently and vice versa; distance is measured in units of percentage recombination (centimorgans, cM).

The overall genetic map lengths in the Graminae are remarkably similar — about 1,500–2,300 cM per genome (see table) — despite a difference of more than ten times in DNA content. Indeed, the relative conservation of genetic map length is seen when distances between closely linked markers are compared across the Graminae^{9,10}.

The genetic distance between loci on an RFLP map often bears little relationship to the physical distance on the chromosome. For instance, α -amylase 2 is the closest to the centromere of a large set of closely linked genes on the short arm of chromosome 1 of barley, and yet *in situ* hybridization experiments place it 70 per cent along the arm towards the telomere¹⁴. This is not to say that genes are absent from the centromere proximal regions (although there is physical evidence to suggest that they are less frequent and that this region is full of repeated DNA¹¹); it may be that they are not associated with polymorphisms and so cannot be mapped, or that they cannot be positioned because recombination is suppressed. Here is where comparative mapping from one species of the Graminae could help position genes in unmappable regions of another.

PETROLOGY

Get out your arc umbrellas

Trevor Falloon

THE power of explosive volcanic eruptions testifies to the importance of volatile substances dissolved in magma. The most abundant volatile, water, may determine a magma's history right from birth: it affects the temperature at which the Earth's mantle melts, the composition of the resulting magma and the potential explosiveness of its eruption. But establishing the volatile content before eruption is far from easy; models of arc magmas (magmas generated along a belt above a subduction zone) have run wild, some having a water content so low as to be negligible, others containing so much water that one reaches for an umbrella. Welcome constraints have now been introduced into the debate by Gaetani and co-workers¹, who set a range of 2 to 4 weight per cent water for parental arc magmas (page 332 of this issue).

Direct measurement of the water content from hot, freshly erupted magma is dangerous and unsatisfactory, because most of the original water is lost through degassing before or during eruption (it has very low solubility in arc magmas at low pressure). The water content of eruptive products (lava, pumice and so on) can also be modified after eruption by secondary processes such as weathering and alteration which may either add or subtract water. But there are several ways of estimating the pre-eruptive contents of arc magmas, and Gaetani and co-workers use one of the more convincing approaches.

What they have done is to study in laboratory experiments the crystallization behaviour of an appropriate composition of parental arc magma under anhydrous and hydrous conditions (see ref. 2 for details of experimental petrology tech-

niques). This approach is powerful because it permits detailed matching of the melt and mineral compositions produced in the experiments with those of natural arc volcanics. Because the difference between the hydrous and anhydrous melts and mineral compositions is so marked, and because their hydrous run products resemble the natural arc volcanics, their conclusion seems inescapable: parental arc magmas must have high and significant water contents (2–4 weight per cent).

Some geologists will consider these figures to be unrealistically high. Parental arc magmas have been widely thought to have low water contents (generally less than 1.5 weight per cent; ref. 3), broadly equal to their K₂O contents⁴, which is consistent with the similar incompatible behaviour of H₂O and potassium (both prefer to enter the melt as opposed to solid phases) during basalt petrogenesis in the sub-oceanic mantle⁵. Supporting this view was the reported water content of quenched submarine glasses dredged from the Mariana arc at a water depth of 1,170 metres⁶, where it was assumed that the pressure of the overlying water column would inhibit degassing. These evolved glasses (2–3.5 weight per cent MgO) have 0.8–1.5 weight per cent H₂O, roughly equal to their K₂O contents of 1–1.4 weight per cent.

Low pre-eruptive water contents have also been considered consistent with the absence of the hydrous silicate mineral amphibole as a phenocryst phase in arc magmas (apart from the more potassium-rich arc magma compositions) and with the abundance of calcium-rich plagioclase phenocrysts, thought to be the low-pressure liquidus phase of hot, anhydrous

magmas. But amphibole stability depends on the chemical composition of the magma (see ref. 7) and its absence is not a proof of low water contents in arc magmas. In addition, as both Gaetani *et al.*¹ and Sisson and Grove⁸ have demonstrated, the compositions of the plagioclase (rich in calcium) and clinopyroxene (rich in calcium, poor in iron) that crystallize from arc magmas are just not right if crystallization took place under relatively anhydrous conditions. This holds regardless of the pressure at which the crystallization occurred. Arc magmas must contain much more water than has been thought.

Additional independent evidence comes from a large collection of quenched submarine glasses from subduction zone settings. In these, Danyushevsky *et al.*⁹ have found that water is substantially enriched relative to K₂O. For example, island-arc tholeiite glasses dredged from water depths of more than 6,000 metres from the Hunter–Matthew ridge in the southern part of the Vanuatu arc have water contents of 3 weight per cent (which Danyushevsky *et al.* consider a minimum estimate of the pre-eruption content for these particular glasses) at a MgO content of 4–4.5 weight per cent and K₂O content of 0.24–0.30 weight per cent. That's a lot of water for a rock series (tholeiitic) that many would have considered relatively anhydrous. These data on dredged submarine glasses are also supported by analyses¹⁰ of melt inclusions hosted in olivine phenocrysts from lava erupted from the Fuego volcano, Guatemala. The inclusions, which represent trapped melt from an evolving magma before eruption, have up to 6 weight per cent H₂O at MgO contents of 4–5 weight per cent (K₂O about 0.8 weight per cent).

So it seems that the experimental petrology approach gives realistic answers when used to determine the pre-eruption water contents of arc magmas. Now what is needed is to incorporate these new constraints into quantitative models of arc magma genesis and into the design of realistic melting and crystallization experiments on arc magmas. □

Trevor Falloon is in the Department of Geology, University of Bristol, Queens Road, Bristol BS8 1RJ, UK.

Circadian clocks à la CREM

Joseph S. Takahashi

DESCARTES argued that the pineal gland is of especial significance as the seat of the soul because it is the only unpaired structure in the human brain. From the molecular view of the rodent pineal provided by Stehle *et al.* on page 314 of this issue¹, it seems that this organ is indeed special. The work centres on the circadian control of gene regulation, and contains some intriguing molecular twists.

The authors have discovered a product of the CREM (cyclic AMP-responsive element modulator) gene that is highly

system. Photic input also reaches the pineal through the same pathway. At night, clock-driven norepinephrine release activates pineal β -adrenergic receptors which stimulate adenylate cyclase and cause an increase in intracellular levels of cAMP (Fig. 2). At the same time, norepinephrine activates α_1 -adrenergic receptors which trigger the phosphoinositide pathway and increase intracellular Ca²⁺. Activated protein kinase C then dramatically potentiates the stimulatory effects of receptors positively coupled to adenylate cyclase. Ultimately this combined β - and α_1 -adrenergic receptor stimulation causes a synergistic increase in cAMP².

Interestingly, cAMP appears to act at three different levels (Fig. 2)³. First, in inducing the transcription of NAT or an NAT regulator; second, in stimulating NAT translation (or that of a regulator); and third, in maintaining NAT in an active form. However, NAT has yet to be purified and analysed at the molecular level, so its regulation cannot yet be directly studied. This is one reason why the new results of Stehle *et al.*¹

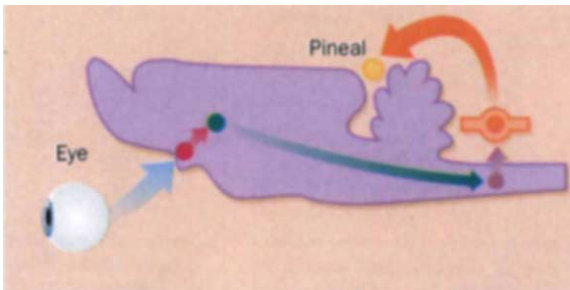


FIG. 1 The neural system regulating pineal *N*-acetyltransferase in the rat. Light is received by the eye and is transmitted by the retinohypothalamic tract (blue) to the suprachiasmatic nucleus (red) which contains a circadian clock. The neural pathway includes the paraventricular nucleus (green), the intermediolateral cell column (purple), the superior cervical ganglion (orange) and the pineal. Adapted from ref. 2.

expressed in neuroendocrine cells, and which encodes a potent repressor of cAMP-induced gene transcription. Levels of messenger RNA of this product — named ICER, for inducible cAMP early repressor — show a remarkable circadian rhythm in the pineal, being high at night and low during the day. Moreover, molecular analysis of ICER shows that it is transcribed from a novel intronic promoter in the CREM gene which preserves the DNA-binding domain but excludes regulatory and activation regions. These results highlight a hitherto unknown level of CREM gene regulation, and suggest that dynamic expression of ICER is a central player in the neuroendocrine axis.

The pineal has long been a model for studying the environmental and circadian regulation of the neuroendocrine system^{2,3}. The mammalian pineal expresses a circadian rhythm in the production of the hormone, melatonin, which is controlled by a circadian fluctuation in the enzyme arylalkylamine *N*-acetyltransferase (NAT). A well-described but tortuous neural pathway governs pineal NAT activity (Fig. 1). A hypothalamic circadian clock in the suprachiasmatic nucleus drives rhythmic noradrenergic input to the pineal by way of the sympathetic nervous

are of such interest.

The CREM gene is a member of the CREB/ATF (CRE-binding/activator transcription factor) family of genes whose products bind the cAMP-responsive element (CRE)⁴. It is notable because it can generate both transcriptional activators and repressors by alternative splicing^{5–7}. Unlike other CREB/ATF members, CREM isoforms are inducible, tissue specific and developmentally regulated^{6,8}. Foulkes *et al.*⁶ have shown that CREM expression in the testes switches during spermatogenesis from antagonist forms (CREM α , β , γ) to an activator form (CREM τ) which has two glutamine-rich regions, Q1 and Q2. The developmental switch in CREM expression during the pachytene spermatocyte stage is regulated by the pituitary hormone, FSH, which causes a high level of CREM τ expression by alternative polyadenylation and an increase in transcript stability⁸. The levels of CREM regulation thus include alternative splicing, alternative translational initiation and alternative polyadenylation.

Stehle *et al.* now document a surprising new level of CREM expression. ICER is transcribed from a second intronic promoter and lacks all exons of CREM 5' to the

1. Gaetani, G. A., Grove, T. L. & Bryan, W. B. *Nature* **365**, 332–334 (1993).
2. Rutherford, M. J. *Eos* **74**, 49–55 (1993).
3. Perfit, M. R. *et al. Chem. Geol.* **30**, 227–256 (1980).
4. Gill, J. B. *Orogenic Andesites and Plate Tectonics* (Springer, Berlin, 1981).
5. Michael, P. J. *Geochim. cosmochim. Acta* **52**, 555–566 (1988).
6. Garcia, M. O. *et al. Geochim. cosmochim. Acta* **43**, 305–312 (1979).
7. Foden, J. D. & Green, D. H. *Contrib. Miner. Petrol.* **109**, 479–493 (1992).
8. Sisson, T. W. & Grove, T. L. *Contrib. Miner. Petrol.* **113**, 143–166; 167–184 (1993).
9. Danyushevsky, L. V. *et al. Earth planet. Sci. Lett.* **117**, 347–362 (1993).
10. Sisson, T. W. & Layne, G. D. *Earth planet. Sci. Lett.* **117**, 619–635 (1993).

γ -exon. An ICER-specific exon between the Q2 region and γ -exon contains the translational initiation site. The ICER gene product, therefore, encodes a 120-amino-acid protein from the carboxy terminus which includes the DNA-binding domains of CREM, but not the amino-terminal phosphorylation domain (P-box) or the two glutamine-rich domains. Like other CREM isoforms, ICER shows

plained by desensitization of receptors or by other steps upstream of cAMP, because it is seen in the presence of cAMP analogues. Instead, it could be that ICER expression, which peaks at a time corresponding to the declining phase of the NAT, leads to repression of NAT/regulator gene expression (Fig. 2). Because ICER itself is induced by cAMP treatment in the pineal, one would im-

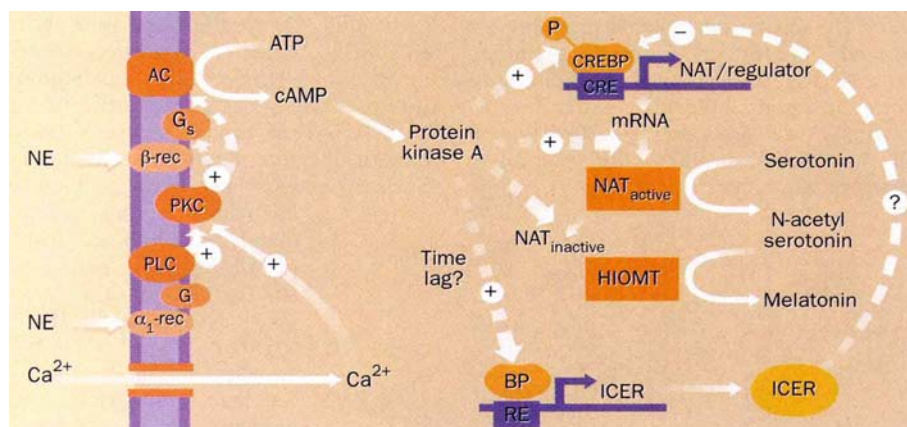


FIG. 2 Proposed scheme for adrenergic regulation of pineal melatonin biosynthesis and ICER gene expression. AC, adenylate cyclase; G, G protein; rec, receptor; NE, norepinephrine; PKC, protein kinase C; PLC, phospholipase C; CREBP, CRE binding protein; HIOMT, hydroxyindole-O-methyltransferase; BP, binding protein; RE response element. Adapted from refs 2, 3.

differential splicing of the two alternative DNA-binding domains and of the γ -exon. But ICER is the first example of a CRE-binding protein that does not have the P-box, which means that the level of expression may be the main determinant of its activity.

In some elegant experiments, Stehle *et al.* show that ICER is a powerful repressor of cAMP-induced transcription (the most potent of the CREM family yet tested). It also appears that ICER is the predominant form of CREM expressed in most tissues, the highest levels being found in neuroendocrine cells such as those of the pineal, pituitary and adrenal glands.

In the rat pineal, ICER is under circadian control. As with NAT, ICER mRNA is induced at night and peaks about 6–9 hours after lights off. The nocturnal elevation of ICER mRNA is under adrenergic regulation because it is blocked by superior cervical ganglionectomy and by β -adrenergic antagonists. Furthermore, treatment with β -adrenergic agonists or with cAMP analogues in organ culture mimics the nocturnal induction of ICER, showing that elevation of cAMP can induce ICER gene expression.

The time course of ICER expression in the pineal both suggests what ICER may do and can explain a long-recognized paradox in the pineal. Continuous treatment of pineal glands with norepinephrine or dibutyryl cAMP causes an increase in NAT which reaches a plateau after 6–12 hours and then decreases to baseline levels⁹. That decrease cannot be ex-

plained by desensitization of receptors or by other steps upstream of cAMP, because it is seen in the presence of cAMP analogues. Instead, it could be that ICER expression, which peaks at a time corresponding to the declining phase of the NAT, leads to repression of NAT/regulator gene expression (Fig. 2). Because ICER itself is induced by cAMP treatment in the pineal, one would im-

agine that ICER should also negatively regulate its own gene expression. If so, the inability to induce ICER expression in the pineal during the day¹ may reflect ICER autorepression. So far, only a few examples of circadian clock control of transcription have been described¹⁰. Along with the *period* gene of *Drosophila*¹¹, ICER now provides an example in which circadian regulation of gene expression is accompanied by a feed-back loop. Perhaps in cases of dynamic gene regulation, simple unidirectional regulation such as the use of an activator may not be adequate. Rather, it may be that we will come to see the importance of negative feedback loops in which both positive and negative regulation of rhythmic gene expression occurs using a 'push-pull' mechanism. □

Joseph S. Takahashi is at the NSF Center for Biological Timing, Department of Neurobiology and Physiology, Northwestern University, Evanston, Illinois 60208, USA.

1. Stehle, J. H. *et al.* *Nature* **365**, 314–320 (1993).
2. Klein, D. C., Schaad, N. L., Nambodiri, M. A. A., Yu, L. & Weller, J. L. *Biochem. Soc. Trans.* **20**, 299–304 (1992).
3. Zatz, M. In *Handbook of Experimental Pharmacology*, Vol. 58/II (eds Kebabian, J. W. & Nathanson, J. A.) 691–710 (Springer, Berlin, 1982).
4. Habener, J. F. *Molec. Endocr.* **4**, 1087–1094 (1990).
5. Foulkes, N. S., Borelli, E. & Sassone-Corsi, P. *Cell* **64**, 739–749 (1991).
6. Foulkes, N. S. *et al.* *Nature* **355**, 80–84 (1992).
7. Laoide, B. M., Foulkes, N. S., Schlotter, F. & Sassone-Corsi, P. *EMBO J.* **12**, 1179–1191 (1993).
8. Foulkes, N. S. *et al.* *Nature* **362**, 264–267 (1993).
9. Klein, D. C. & Weller, J. L. *J. Pharmac. exp. Ther.* **186**, 516–527 (1973).
10. Takahashi, J. S. *Curr. Opin. Genet. Devel.* **3**, 301 (1993).
11. Hardin, P. E. *et al.* *Nature* **343**, 536–540 (1990).

DAEDALUS

Traffic pressure

MANY racing cars are designed to have 'negative lift'. The airflow over them pushes them down on the road, increasing the traction of their tyres. Daedalus is now taking this idea to extremes. His new 'suckercar' has a massive extractor fan to suck air from underneath it. The force holding it down on the road can exceed its own weight. So it can adhere to a vertical wall, and can even run upside-down across a ceiling.

The design is largely a direct inversion of conventional hovercraft practice. The fan of the suckercar consumes very little power; a mere 0.02 atmospheres of underpressure can exert the required force, and a skirt brushing the road all round the vehicle restricts inward leakage of air. The fan, however, is continuously variable. At every instant, it delivers the exact adhesive suction required.

On the open road, the fan runs quite slowly, giving just enough adhesion to hold the car's smooth, energy-efficient tyres firmly down on the tarmac. During braking, acceleration or cornering, it sucks harder to prevent skidding. In an emergency, it can suck the car down so ferociously as practically to anchor it to the road. This powerful adhesion could be vital in a crash. A firmly anchored suckercar could not be overturned or kicked along by a collision, and its occupants would be safe from whiplash and other accelerational injuries.

On a wet or icy road, quite vigorous suction is needed to maintain traction. But in serious flooding or in deep soft snow, the fan can be reversed. It then blows downwards, lifting the vehicle and permitting it to make slow but safe progress as a true hovercraft.

The suckercar comes into its own on vertical or inverted surfaces. The full power of the fan is needed to hold the vehicle in place. But it can then navigate the most forbidding mountainsides and overhanging cliffs. Elevated motorways will suddenly double in capacity as their undersides become available; sucker traffic will crowd the hitherto wasted walls and roofs of road tunnels. Sucker fire engines and ambulances will be able to drive straight up the sides of stricken tower blocks, park safely under full suction, rescue people from the top floors, and drive down again.

Weak or pot-holed roads may be a problem. The suckercar's 'negative air-cushion' should be too weak to inhale manhole covers or loose cobbles; but open drains and deep pot-holes may gust air unpredictably into it. Daedalus is fitting fast-acting servos linked to a constantly pumped reserve vacuum tank. For a crucial moment, it can suck as hard as such cavities can blow. David Jones

Viral burden in AIDS

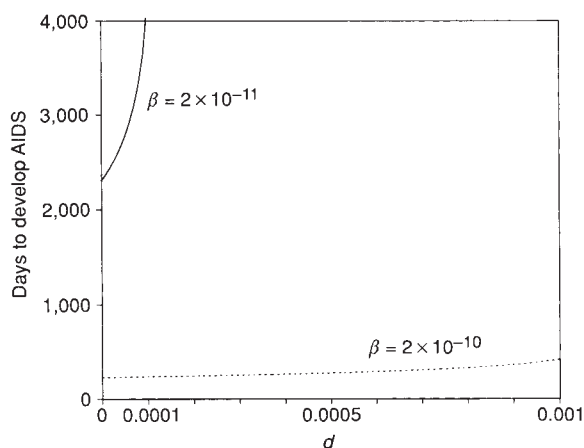
SIR — Recent criticism¹ of studies of HIV infection in lymphoid tissues^{2,3} claim that HIV cannot cause AIDS by a direct virus-mediated cytopathology. Sheppard *et al.*¹ state that a frequency of 1% actively infected lymph node cells is “minuscule compared to the regenerative capacity of the immune system” and that low prevalence of infected cells is “the central paradox of human immunodeficiency virus (HIV-1) infection”. We show here how some simple calculations can resolve this apparent paradox.

The myth of the massive regenerative power of the human immune system seems to originate in extrapolation from mouse data⁴. Regeneration of human immune systems is probably not so rapid. Patients whose peripheral T-cell counts are halved by radiation therapy take 3–4 years to regain a normal count^{5,6}. Such patients have equilibrium counts of around 2,500 cells per mm³, and at 1,000 days post-radiotherapy their counts are around 2,000 cells per mm³. A ‘back-of-the-envelope’ calculation based on these observations yields a rough estimate for the net death rate of T lymphocytes of 1

per 1,000 cells per day. (Net death rate is the difference between the proliferation rate and the death rate.) A very different estimate for this parameter (1 per 10,000 cells per day) is implied by information on the survival of lymphocytes with stable chromosome damage in the years following radiotherapy⁵. These cells pass their damage to one daughter cell on division, so the rate at which they disappear estimates their gross death rate, and hence yields an upper bound for their net death rate.

These parameters can be used in a rough calculation of the time for HIV infection to reduce the CD4⁺ lymphocyte count to one-tenth of the uninfected level. The figure illustrates the outcome of such calculations, showing that if the net death rate of lymphocytes is around 1 per 1,000 cells per day, then all that is surprising is that with 1% of cells infected it takes so long for AIDS to develop. If, on the other hand, the net death rate is tenfold smaller, then a reduction of the population to one-tenth of its original size over about 10 years is understandable. All that is surprising is that there should be such a small difference between the proliferation rate and death rate of human peripheral CD4⁺ cells.

The calculations illustrated in the figure contain a free parameter, β , the cell-to-cell infection rate. The calculations were performed assuming that 1 in 100 cells are actively infected during the asymptomatic phase and choosing a value of β close to that which maximizes the number of days to develop AIDS. If further technological advances show that 1 in 10 cells are actively infected, exactly the same picture could be constructed by reducing the values of β by a tenth. The proportion of cells actively infected and the parameter β never appear separately. Thus to speak of a paucity of infected cells in the complete absence of quantitative information on the cell-to-cell infection rate is meaningless. In the same way, it is difficult to predict the length of the AIDS incubation period without reliable information on the normal population dynamics (and particularly the net death rate) of CD4⁺ cells.



AIDS incubation period as a function of the net death rate (d) of uninfected CD4⁺ cells. Suppose that early in infection 1% of cells are actively infected and that the number of infected cells remains roughly constant during the asymptomatic phase. Suppose also that the death rate of uninfected cells is augmented by the presence of actively infected cells so that the immigration–death process describing the dynamics of uninfected cells becomes $dX/dt = \Lambda - dX - \beta XY$, where X is uninfected CD4⁺ cells, Y is infected CD4⁺ cells and Λ is cell immigration rate. The time taken until the CD4⁺ count falls to one-tenth of its pre-infection value is:

$$T = \frac{1}{d + \beta Y} \ln \left[\frac{10\beta Y}{\beta Y - 9d} \right]$$

Clearly, this is defined only for $\beta Y > 9d$, and will be largest for values of β just above this value. The two lines are drawn with $\beta Y = 10d$, assuming that $Y = 5 \times 10^7$, that is, that an uninfected person has a CD4⁺ count of 5×10^9 cells in total and that early in infection 1% of those are actively infected.

Populations (of cells or of animals) in which the proliferation rate is close to the death rate are particularly prone to having their density suppressed by low-prevalence infections⁷. Furthermore, as outlined above, if the net death rate is very small, one would expect such depression in population size to take place rather slowly. Thus, with certain assumptions about parameters about which very little is known, the regenerative power of the immune system, the low prevalence of HIV-infected cells, and the long time from infection to illness with AIDS form a coherent whole.

Angela McLean

Zoology Department,
University of Oxford,
South Parks Road,
Oxford OX1 3PS, UK

Colin Michie

Department of Paediatrics
and Neonatal Medicine,
Royal Postgraduate Medical School,
London W12 0NN, UK

1. Sheppard, H. W., Ascher, M. S. & Krowka, J. F. *Nature* **364**, 291 (1993).
2. Pantaleo, G. *et al.* *Nature* **362**, 355–358 (1993).
3. Embretson, J. *et al.* *Nature* **362**, 359–362 (1993).
4. Sprent, J. in *Handbook of T Cells in Immune Recognition* (eds Loo, F. & Roelants, G. E.) 59–82 (Wiley, New York, 1978).
5. Buckton, K. E. *et al.* *Nature* **214**, 470–474 (1967).
6. Michie, C. A. thesis (Univ. London, 1993).
7. Anderson, R. M. & May, R. M. *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, 1991).

SIR — Sheppard *et al.*¹ question work^{2–6} supporting “a simple cytopathic model for AIDS pathogenesis”. The arguments of Sheppard *et al.* are similar to those of other HIV iconoclasts who insist that too little HIV and too few HIV-infected cells are present to have a serious detrimental impact. It is ironic that some of the same authors⁷ who dismiss the “AIDS–drug” hypothesis of Peter Duesberg appear to espouse this so-called “central paradox”, a main tenet of Duesberg’s case that HIV is not the cause of AIDS⁸. Like Duesberg, their arguments are based on quantitative interpretations of a single early attempt⁹ to detect HIV-infected cells by *in situ* hybridization, an insensitive technique that cannot be considered quantitative, and which has since been superseded by more sensitive, quantitative assays for HIV-infected cells. Sheppard *et al.* also fail to consider work on HIV cyto-

Scientific Correspondence

Scientific Correspondence is a relatively informal section of *Nature* in which matters of general scientific interest, not necessarily those arising from papers appearing in *Nature*, are published. Because there is space to print only a small proportion of the letters received, priority is usually given according to general interest and topicality, to contributions of fewer than 500 words, and to contributions using simple language.

pathogenesis indicating that direct cytopathic effects of HIV are adequate to explain the CD4⁺ T-cell depletion which is the hallmark of AIDS.

Sheppard *et al.*¹ state that "In a cytopathic model, the rate of CD4⁺ cell loss would be determined by the number of actively infected cells, and even a frequency of one per 100 lymph node cells (based on 1/10 infected and 90 per cent latent) is still minuscule compared to the regenerative capacity of the immune system". But the regenerative capacity of the human immune system has not been measured, particularly in AIDS patients. Components of the immune system, such as the lymph nodes, may not have infinite capacity for self-renewal. When HIV kills cells responsive to certain antigens there may not be mechanisms to restore them or to re-educate other cells to respond to these antigens. Furthermore, Sheppard *et al.* grossly underestimate the effects of acute HIV cytopathology on CD4⁺ cell depletion. In cell culture experiments, for every cell persistently infected with HIV there are typically 10–20 cells killed^{10,11}. *In vivo*, most CD4⁺ cells dying after HIV infection will be rapidly engulfed by scavenger cells such as macrophages. Thus, *in vivo* (as opposed to *in vitro*) the actively infected cells at any cross-section in time represent only a fraction of cells dying from the acute cytopathic effects of HIV. Moreover, Sheppard *et al.* conclude that active HIV infection is required to compromise cell viability, but this is not supported by available data. Cells persistently infected by HIV do not have the same viability as uninfected cells. Typically, HIV-infected T-lymphoblastoid cells are only 95% viable whereas uninfected cells are greater than 99% viable¹⁰.

The second point made by Sheppard *et al.* is that the high levels of plasma virus observed by Piatak *et al.*⁴ were about 99.9% non-culturable, which Sheppard *et al.* believe does not support the cytopathic model. But the fact that most HIV virions present in the circulation no longer seem

to be infectious is not an argument against the HIV hypothesis. HIV is not known as a stable virus, so it comes as no surprise that most plasma virus is not culturable. However, in the microenvironment of the lymph node or elsewhere, newly released and intact HIV could do substantive damage before entering the circulation. Moreover, direct toxic and immunosuppressive effects of inactivated HIV or HIV proteins are well documented^{11–14}. In fact, there may be far more virion equivalents of HIV proteins present in the circulation than genomic RNA. It is not unusual to find thousands of picograms of the major HIV capsid protein p24 per ml of serum and elsewhere in HIV-infected people. Dividing Avogadro's number by the molecular mass of p24 and assuming approximately 1,000 p24 molecules per HIV virion gives an estimate of 2.5×10^4 particles per pg p24. The highest number of virion RNA equivalents of HIV measured by Piatak *et al.* (1.09×10^7) is an order of magnitude lower than estimates based on the 5,406 pg p24 protein per ml present in this patient (1.35×10^8 virion equivalents).

Finally, the recent data^{2–4} may indeed explain the relative ineffectiveness of AIDS therapy with nucleoside analogues, but for reasons opposite to those sug-

gested by Sheppard *et al.* Their suggestion that the ineffectiveness of AZT and related drugs in treatment of HIV disease is due to the slow dissemination of the infected cell burden is contrary to the fact that even early after infection there is already a massive infection of the lymph nodes and elsewhere^{2,3}. It has been known for years that AZT does not affect HIV production from already infected cells (except as a result of drug cytotoxicity), explaining the high levels of plasma virus even in the presence of the drugs. Furthermore, nucleoside analogues that block the reverse transcriptase step do not block early interactions of HIV with CD4⁺ cells. These early membrane interactions can be deleterious to cells even if further steps in HIV replication are blocked^{12,15}. Thus, inactivated or defective HIV or structural proteins of the virus released from cells infected before drug therapy can be cytotoxic even in the presence of nucleoside analogues.

Robert F. Garry

Department of Microbiology and Immunology,

Cesar D. Fermin

Department of Pathology,
Tulane University School of Medicine,
New Orleans,
Louisiana 70112, USA

Which source of mercury pollution?

SIR — Nriagu¹ concludes that the Spanish American silver mines were partly responsible for the high global background concentrations of mercury now being reported. Although there is no doubt that the atmosphere is the main pathway for the global distribution of heavy metals, I disagree with Nriagu's conclusion.

The total natural input of mercury to the atmosphere ranges from 25,000 to 50,000 tonnes per yr (ref. 2), the main source being continental degassing. In addition, the contemporary anthropogenic input of mercury to the atmosphere falls in the average range of 3,560 to 13,000 tonnes per yr (refs 2, 3), the main source being burning of fossil fuels. The residence time of mercury in the atmosphere has been estimated as 5.7 yr (ref. 2). After atmospheric transport, most of the mercury from natural and anthropogenic inputs goes to the oceans through precipitation and fluvial transport. Thus, the open oceans are the primary global reservoirs and accumulators of mercury. The average concentration of mercury in unpolluted sea water has been reported to be 0.1 p.p.b. (refs 4, 5), the total amount of mercury in the oceans ranging from 41 million to 135 million tonnes^{4–7}. This is about 200–700 times Nriagu's estimate of the cumulative loss of mercury in Central and South America between 1570 and 1900. And because mercury binds tightly to many organic and inorganic materials,

a fraction of that total amount settles to the bottom of the ocean every year.

Thus the contribution of the Spanish American silver mines to the high background concentrations of mercury in the global environment is negligible. Moreover, owing to the relatively long time since the mines were operational, it is probable that most of the mercury lost during the refining of silver via the *patio* process is now in the open oceans. After all, the cumulative input of mercury to the atmosphere from the combustion of fossil fuels since global industrialization is surely much higher than that from the refining of silver and gold so far (including the gold rush in California in the past century and the present gold rush in Amazonia).

Julio A. Camargo

Institute of Wildlife and Environmental Toxicology,
Clemson University,
Pendleton,
PO Box 709,
South Carolina 29670, USA

1. Sheppard, H. W., Ascher, M. S. & Krowka, J. F. *Nature* **364**, 291 (1993).
2. Pantaleo, G. *et al.* *Nature* **362**, 355–358 (1993).
3. Embretson, J. *et al.* *Nature* **362**, 359–362 (1993).
4. Piatak, M., Saag, M. S. & Yang, L. C. *Science* **259**, 1749–1754 (1993).
5. Temin, H. M. & Bolognesi, D. *Nature* **362**, 292–293 (1993).
6. Cohen, J. *Science* **260**, 292–293 (1993).
7. Ascher, M. S., Sheppard, H. W., Winkelstein, W. Jr & Vittinghoff, E. *Nature* **362**, 103–104 (1993).
8. Duesberg, P. H. *Pharmac. Ther.* **55**, 201–277 (1992).
9. Harper, M. E., Marselle, L. M., Gallo, R. C. & Wong-Staal, F. *Proc. natn. Acad. Sci. U.S.A.* **83**, 772–776 (1993).
10. Rasheed, S., Gottlieb, A. A. & Garry, R. F. *Virology* **154**, 395–400 (1986).
11. Volsky, D. J., Zeira, M. & Loyter, A. In *Membrane Interactions of HIV* (eds Aloia, R. C. & Curtin, C. C.) 167–186 (Wiley-Liss, New York, 1992).
12. Grewe, C., Beck, A. & Gelderblom, H. R. J. *AIDS* **3**, 965–974 (1990).
13. Henderson, L. A., Qureshi, N. M., Rasheed, S. & Garry, R. F. *Clin. Immun. Immunopath.* **48**, 174–186 (1988).
14. Miller, M. A., Garry, R. F., Jaynes, J. M. & Montelaro, R. C. *AIDS Res. Hum. Retrovir.* **6**, 511–519 (1991).
15. Fermin, C. D. & Garry, R. F. *Virology* **191**, 941–946 (1992).

1. Nriagu, J. O. *Nature* **363**, 589 (1993).
2. US Environmental Protection Agency *Ambient Water Quality Criteria for Mercury* (EPA, Washington, DC, 1980).
3. Nriagu, J. O. & Pacyna, J. M. *Nature* **333**, 134–139 (1988).
4. Hammond, A. L. *Science* **171**, 788–789 (1971).
5. OECD *Mercury and the Environment* (OECD, Paris, 1974).
6. D'Itri, P. A. & D'Itri, F. M. *Mercury Contamination* (Wiley, New York, 1977).
7. US National Academy of Sciences *An Assessment of Mercury in the Environment* (NAS, Washington, DC, 1978).

Universal foodstuff

Richard Davenport-Hines

For God, Country and Coca-Cola: The Unauthorised History of the World's Most Popular Drink and the Company That Makes It. By Mark Pendergrast. *Weidenfeld and Nicolson*: 1993. Pp. 556. £20, \$27.50.

MARK Pendergrast writes in a breathless, enthusiastic prose that would be easy to tease. He has a Coca-Cola-centric view of world history that sometimes surpasses belief. Every great event is interpreted in terms of the world's thirst for this fizzy brown drink which is 99 per cent sugar-water with caramel colouring: "The Japanese didn't realize that by bombing Pearl Harbor, they were indirectly giving the Coca-Cola Company a worldwide boost that would ensure . . . global dominance of the industry", he writes of US entry into the Second World War. "It's likely that the Japanese weren't thinking about soft drinks at all, although four Hawaiian Coke coolers were in fact martyred that day." Later on, the US civil-rights movement is reduced to the same level: "By demanding equal rights to Coca-Cola, the civil-rights activists were striking at the heart of Southern and American culture." Yet despite these recurrent absurdities, an excessive attention to detail and a tight chronological structure which means that trivia constantly interrupt the essential narrative, there are virtues in the book.

Pendergrast is a business journalist who grew up in Coca-Cola's home town of Atlanta, Georgia, on a street known locally as Coca-Cola Row. He has had access to the company's archives and interviewed a multitude of people associated with it, some of whom express themselves with surprising profanity. There are some good anecdotes, and many fascinating side-lights on the trashier side of American culture. The drink is traced from its origins in the 1880s as a cocaine-laced panacea peddled by quacks for "neurasthenics" to a twentieth-century mass-distributed cultural icon. Nowadays the multinational company has annual sales of \$13 billion: its chief executive was paid a salary in 1991 of \$86 million.

For medical historians there are several points of interest. The first is the early marketing of the drink as a "nerve tonic" for what Pendergrast calls the "turbulent,

inventive, noisy, neurotic new America" of the 1880s, and his description of the advertising effort that went into pushing patent medicines. Conflicting and often inaccurate perceptions of cocaine are also described. Its cocaine content was tolerated when Coca-Cola was "a soda fountain drink for upper-class urban white professionals", but around 1900 scare stories of "Negro coke fiends" attacking



Beatle-juice — pop advertises pop.

their bosses or raping white women led to agitation against the beverage. In 1903 the cocaine ingredient was removed by the manufacturers, who in later years repeatedly denied under oath that the drink had ever contained cocaine. This official denial has continued into the 1990s.

The marketing of Coke typified what George Orwell called "Admass": the mix of mass culture with advertising in an appeal to the lowest common denominator. The company's annual budget for advertising and marketing now runs at \$4 billion. "We're selling smoke", the creative staff of their advertising were always reminded. "They're drinking the image, not the product." Initially people drank it because it seemed naughty or even dangerous (even after the cocaine in the original formula was dropped). Advertising played on this. Pendergrast reproduces a 1907 promotional poster featuring an exhausted beauty, lying on a tiger-skin rug with her skirt rucked above her stockings and an empty Coca-Cola bottle beside her, with the advertising caption

"Satisfied". More recently, there has been a preference for a more clean-cut, family-orientated image, which is why the drink's cocaine-based origins are so forcibly suppressed by the company.

The turning point in the Coke story was in 1942. Coca-Cola had recently hired as head of its export department a man who had been campaign manager to President Franklin Roosevelt. He persuaded the US government to exempt the company from rationing, and, just as astonishingly, to let Coca-Cola build state-subsidized bottling plants wherever US troops were fighting. Coke thus acquired plants on every continent except Antarctica and an advantage that Pepsi-Cola has never been able to approach.

Coke's greatest reverse occurred in 1985. The company was alarmed by Pepsi-Cola making inroads of one per cent into its market share, and therefore changed its drink's flavour, marketing it as "New Coke". Consumers were revolted, not by the change of taste, but by the tampering with an icon. Coke for them had to represent permanence, stability and universality. "Changing Coke is like breaking the American Dream", complained one customer. "There are only two things in my life, God and Coca-Cola", wrote another: "Now you have taken one of these things away from me". Such emotional investment in a beverage does not fill Pendergrast with despair. This failed marketing experiment was duly abandoned.

Noting that "the world's first coin-operated vending machine, invented in the first century AD, dispensed holy water", Pendergrast compares Coca-Cola to "a new religion", and approvingly quotes an employee declaring that his job "is like being a representative from the Vatican, like you've touched God". Pendergrast continues: "The Coca-Cola religion has no morality, no commandment other than increased consumption of its drink". For him, "therein lies the true beauty of capitalism at its best". Given the identification of Coca-Cola with the American way of life, this book unwittingly gives a fillip to anti-American feeling by such naivety, cultural mediocrity and arrogance. It is though a case study of what is possibly the world's greatest marketing success, and gives a clear image of multinational corporate culture. □

Richard Davenport-Hines is at 51 Elsham Road, London W14 8HD, UK.

Manipulating mankind

John Habgood

Perilous Knowledge: The Human Genome Project and Its Implications.

By Tom Wilkie. *Faber and Faber*: 1993. Pp. 195. £14.95.

The Human Body Shop: The Engineering and Marketing of Life. By Andrew Kimbrell. *Harper San Francisco*: 1993. Pp. 348. \$22, £8.99.

THE Human Genome Project must be unique among major scientific programmes in having three per cent of its budget set aside for the study of its ethical implications. In a three-billion-dollar budget, that is a lot of money. It is a measure of the concern felt by some of the initiators of the project about its potential social impact.

Why is the project being done at all? In part it reflects the growth in the relative importance of genetic disease as other potentially fatal diseases have been eliminated. Thalassaemia, for instance, manifested itself in Cyprus only as malaria deaths declined. Given that there are some 4,000 diseases known or suspected to be due to single-gene mutations, a reliable and comprehensive genetic map could be an enormous aid to diagnosis, even though the problems of effective treatment are likely to remain formidable.

In part the project also represents a successful bid to bring biology into the 'big science' league, not least because the US weapons programme was looking for ways to diversify into the study of the genetic effects of radiation. The motive that finally prevailed, however, seems to have been the sheer challenge of unravelling such a massively complex structure, of finding, as some put it, "the holy grail of human biology". Time has tended to change all that. There is more emphasis now on mapping and elucidating the structure of identifiable genes, rather than trying to sequence the whole genome, a large proportion of which seems to have no clear function. And motives have become more mixed, as evidenced by frustrating disagreements over whether or not sections of the genome can be patented.

Commercial interests also loom large in some of the ethical issues. Knowledge is not always an unmixed blessing. Given that it may be possible to diagnose a genetic predisposition towards a wide variety of conditions from heart disease to schizophrenia, what are the implications for health care, for instance, and how will it be possible to avoid discrimination in employment? These problems are already with us, and at a conference last year in the

United States a main topic of conversation was whether a privately funded health system can any longer be operated without gross unfairness.

Bigger problems about the effective treatment of genetic conditions, and the potential for making genetic 'improvements' to individuals or their progeny, belong mainly to the future. But they will certainly arise. Tom Wilkie foresees a new generation of "snake oil salesmen" peddling illegitimate claims about what the new biotechnology can do for the desperate or the gullible.

Previous experience of the use of this kind of knowledge is not encouraging. Lessons have been learnt from the mishandling of the US screening programme for sickle-cell anaemia, which resulted in widespread discrimination against many who were at no risk of contracting the disease. But what, for example, are the ethics of screening for conditions where there is no effective cure? The commercial exploitation of genetically engineered human growth hormone is another unsavoury precedent. This was a product looking for a market, and finding one by trying to alter the definition of normal human height.

Wilkie writes attractively, forcefully and unsensationally. For those who want an accurate but not too technical account of what is happening in genetics, and the ethical problems to which the discipline gives rise, *Perilous Knowledge* could scarcely be bettered. He points beyond the immediate issues, which are difficult enough in all conscience, to the fundamental change in perspective which this completion of Darwin's work may produce. Perhaps the most destructive consequence may be the reinforcement of reductionism, the sense that humanity is now 'understood' in a newtonian sense, and hence the false idea that human beings are simply to be identified with their genes, and no more. He wrestles with the concept of human uniqueness and while, unnecessarily in my view, rejecting a religious interpretation of this, ends up with a high view of the moral worth and transcendence of human life.

The Human Body Shop paints an alarming picture of what could happen, and in many respects is already happening, when this sense of the transcendent significance of human life is ignored, and when the human body is made subject to commercial pressures. The book is much more balanced and sober than its title and a somewhat over-the-top foreword by Jeremy Rifkin might suggest.

Andrew Kimbrell is an American lawyer who in recent years has made a detailed study of the social impact of biotechnology. He surveys an enormously wide field from the sale of blood to the collecting of fetal parts, from transplant surgery to surrogate motherhood, and

from the search for a perfect baby to the patenting of life. His style is anecdotal. Much of his material is drawn from court cases in the United States. The use he makes of newspaper articles may arouse some doubts. But even if only half of what he records is accurate, it remains a powerful indictment of attitudes and practices that tend to treat the human body as a commodity, ripe for being serviced by suppliers of spare parts.

A single illustration must suffice. In June 1991, a 14-month-baby was used to provide bone marrow for her 18-year-old sister who was suffering from leukaemia. Their parents subsequently announced that they had conceived the child solely for the purpose of providing compatible bone marrow for their other daughter. In this instance there was a happy ending. But suppose prenatal diagnosis had revealed that the bone marrow would not be compatible. Would they have had an abortion and tried again?

Philosophies that treat the body as a machine and the market as the unquestioned supplier of human wants, are endemic to the whole of the Western world. As manipulative techniques become more powerful, the pressures to use them, initially no doubt in response to urgent need and eventually to feed human vanity and selfishness, could become irresistible. Choices need to be made now about the direction in which we wish our civilization to move. □

John Habgood, Archbishop of York, is at Bishopthorpe Palace, Bishopthorpe, York YO2 1QE, UK.

New in paperback

Saltmarsh Ecology by Paul Adam. Cambridge University Press, £19.95, \$37.95.

The Gods of the Celts by Miranda Green. A "sourcebook on the Celts and their religion for the historian, archaeologist and general reader". Alan Sutton, £9.99.

Quantum Processes in Semiconductors by B. K. Ridley (3rd edn). Oxford University Press, £17.50.

Ecology, Economics, Ethics: The Broken Circle edited by F. Herbert Borman and Stephen R. Kellert. Yale University Press, \$14.95, £10.

Documentary Archaeology in the New World edited by Mary C. Beaudry. CUP, £14.95, \$19.95.

The Molecular and Cellular Biology of the Yeast *Saccharomyces*. Volume 2: Gene Expression edited by E. W. Jones, J. R. Pringle and J. R. Broach. Cold Spring Harbor Laboratory Press, \$65.

Humoral Factors edited by E. Sim. IRL/OUP, £27.50.

Essential Developmental Biology: A Practical Approach edited by C. D. Stern and P. W. H. Holland. IRL/OUP, £25.

Animal houses

James Rachels

Zoos and Animal Rights: The Ethics of Keeping Animals. By S. St C. Bostock. Routledge: 1993. Pp. 227. £35, \$59.95 (hbk); £10.99, \$15.95 (pbk).

THE debate about animal rights is nothing if not strident. Often it seems to be a shouting-match between activists who don't know very much about the practices they are criticizing and animal-users who are so intent on vindicating themselves that ethical questions cease to matter. *Zoos and Animal Rights* is a contribution to the debate of a much higher order. The author of this engaging book is himself a member of the profession whose practices are being scrutinized: he is the education officer at Glasgow Zoo in Scotland. But Bostock also has a doctorate in philosophy and possesses a keen understanding of the ethical problems that beset his line of work. Is zoo-keeping compatible with a proper regard for the rights and interests of the animals on display? In *Zoos and Animal Rights* this question gets a thorough and intelligent airing.

Of course, one would not expect Bostock to conclude that his job should be abolished, and he does not. He defends zoos on a number of grounds, as resources or conservation, education and science. He also argues that we should not object, on the grounds of animal welfare, to keeping animals in zoos if they would be worse off in the wild, and he describes in detail how various animals fare in both places. Along the way he provides a wealth of information about the history of zoos and about the hows and whys of contemporary animal-keeping. This background material alone makes *Zoos and Animal Rights* worth reading, even if one has no interest in the moral debate.

Bostock does not defend zoos by contending that animals have no rights. On the contrary, he ascribes extensive rights to them and argues that, in order to be morally defensible, zoos must operate compatibly with these rights. To his considerable credit he does not gloss over the suffering that zoo animals are sometimes made to endure. He knows, better than most of us, that many zoos are awful places. It is no wonder. Providing an appropriate environment for exotic animals requires not only a concern for their well-being but a great deal of esoteric

knowledge and considerable resources. Thus many zoos that seem like pleasant places to take the children on a Sunday afternoon are nevertheless intolerable from the point of view of animal welfare. But Bostock argues that it is unfair to condemn all zoos indiscriminately. There are bad zoos, he says, but there are good ones too.

The more radical critics are not likely to be placated, however, not only because conditions in bad zoos can be so appalling (and the bad zoos greatly outnumber the good ones) but also because the treatment that Bostock describes of animals even in good zoos may be deemed unacceptable. In his honest depiction of the history and practice of animal-keeping, he gives the critics a lot of ammunition. Moreover, despite his evident humanitarianism, Bostock is willing to tolerate practices that others find repugnant. He candidly notes that killing "food animals" to be con-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

German engraving (c. 1870) of an Ancient Greek zoo.

sumed by carnivores is "regrettable but unavoidable". Others will conclude that it is better not to keep animals at all than to keep some at the cost of killing others. (Of course, this reasoning also implies that we should be vegetarians, as killing "food animals" for our own consumption is no different. But supporters of animal rights will have no trouble with such a conclusion.)

Nevertheless, if stridency can be avoided, people on both sides of the animal-rights debate can learn a great deal from *Zoos and Animal Rights*. Bostock has written a very good book, full of interesting information and moral argument that deserves to be taken seriously. It is too bad that there is no book about the use of animals in medical research that has the same virtues. If there were, it would be invaluable. □

James Rachels is in the Department of Philosophy, University of Alabama at Birmingham, Birmingham, Alabama 35294-1260, USA.

Discovery and prevention

Vitaliy I. Goldanskii and
Stanislav N. Rodionov

In the Shadow of the Bomb: Physics and Arms Control. By Sidney D. Drell. AIP: 1993. Pp. 358. \$30, £25. (Distributed outside the United Kingdom by Oxford University Press.)

SIDNEY Drell, the distinguished American scientist and public figure, describes his life as divided "between pursuing the dream of discovery and working to avoid the nightmare of a nuclear holocaust". He is known as a first-class high-energy physicist and a former deputy director of the Stanford Linear Accelerator Center and president of the American Physical Society. Since the early 1960s he has been actively involved in national security, working for the US President's Science Advisory Committee, the Arms Control and Disarmament Agency, the National Security Council and currently as a key member of the Stanford Center of International Security and Arms Control. His work on security issues has had a worldwide influence: President Mikhail Gorbachev held it in particular high regard.

All these various strands of Drell's career are represented in *In the Shadow of the Bomb*, a collection of articles written between 1974 and 1992, most of them in the 1980s. Because the book is intended for a wide audience, the three articles on elementary particles assume a semipopular style like that of *Scientific American*. Nevertheless, they succeed in providing a good description of the main results and problems in high-energy physics, and are splendid examples of how to make complicated science accessible to the layperson.

Here too are Drell's testimonials to his friends and colleagues (such as Wolfgang Panofsky and Victor Weisskopf), many of whom share his scientific, political and public views, as well as his passion for music. There is also a section devoted to Andrei Sakharov, a close friend and a fellow "insider in two spheres". There are many scientists today — and not only in the United States — who were responsible for the creation of weapons of mass destruction and who are now fighting for their total elimination. Sakharov was the archetypal nuclear scientist turned peace campaigner; his admirers will appreciate Drell's descriptions of his contributions to science, human rights and arms control.

Two articles reprinted in this book deserve special attention because they played a peculiar role in the vicissitudes of Sakharov's life. When "The Impact of a Public Constituency" and "The Moral

Issue and Deterrence" first appeared, Sakharov was in exile in Gorki; his response to these articles, which was published in the June 1983 issue of *Foreign Affairs*, provoked the wrath of Soviet and Communist Party officials. Sakharov was denounced as a hypocrite, accused of being a warmonger who had betrayed his motherland, and was contrasted unfavourably with "peace-defender" Drell. The most notorious article attacking Sakharov (and ostensibly in support of Drell) appeared in *Izvestia* on 3 July 1983 under the title "When Honour and Conscience are Lost" and was signed by four Soviet academicians. It so happened that one of us (V. I. G.) was constantly urged to go to the Communist Party Central Committee and provide the fifth signature. It was not easy to resist this persistent onslaught. As Sakharov writes in his *Memoirs* (Knopf, 1990), the intended fifth author "had refused [to be a signatory] saying that he'd need to get a look at [Sakharov's] article first. . . . He was told that there wasn't time for that, and so he escaped with his honor intact."

Nearly two-thirds of Drell's book cov-

ers his work on arms control, a subject that he considers to be "an important part of our national security". For the past 15 years, Drell has made an independent and original examination of many arms-control problems, including US-Soviet strategic balance, nuclear deterrence, the Strategic Defense Initiative ('Star Wars') and the testing and safety of nuclear warheads. It is his belief that "our scientific community as a whole has a special obligation to assist society to understand the implications of the products of our scientific advances and to shape the applications of those products in ways more beneficial than dangerous to the human condition". Proponents of this admirable conceit will be among the most enthusiastic readers of this wide-ranging and informative book. □

Vitaliy I. Goldanskii is in the Institute of Chemical Physics, Russian Academy of Sciences, 4 Ulitsa Kosygina, Moscow 117334, Russia; Stanislav N. Rodionov is in the Space Research Institute, Russian Academy of Sciences, 84/32 Profsoyznaya Street, Moscow 117810, Russia.

Images of cellular mechanisms

Harold P. Erickson

The Machinery of Life. By David S. Goodsell. Springer: 1993. Pp. 132. £25.50, \$29.

THE goal of this book is straightforward: to present scale model images of the machinery of life, including its complex macromolecules, substrates and large assemblies. The actual achievement is far grander — the drawings are among the most instructive one can find in structural biology, and the mechanisms of life are elegantly explained in only 132 pages.

The genius of the book is in its simplicity. Drawings of molecules are shown at only three magnifications: 1, 10 and 30 million-fold. Structures are presented in only two styles. At the lowest magnification, protein molecules are drawn as simple outlines to give a rough idea of their size and shape: ribosomes have distinctive beaks, antibodies are Y-shaped, DNA is a rope with the pitch correctly indicated. At the higher magnifications, molecules are presented as space-filling atomic models. The consistency of this presentation makes it easy to relate the molecules in the different figures. Most of the drawings are black and white, but there are 16 beautiful colour plates of representative molecules, using red and blue to show charged atoms, green or yellow for other atomic features. The comparison of molecules is very instructive, and the relative dimensions of ATP,

enzymes, DNA and ribosomes quickly become second nature.

Probably the most novel and important perspective in these drawings is to show groups of molecules at their actual concentration *in vivo*. That the bacterial cytoplasm is "a teeming soup of molecules" is vividly emphasized by drawings showing the jumble of mitochondria, enzymes and RNA molecules that fill 30 per cent of the cytoplasmic volume. The tangled mixture of double- and single-stranded DNA, histones and polymerases in the nuclear region is daunting. A few minutes study may be needed to overcome the feeling of claustrophobia, but will reward the reader with a lasting impression of the density and complexity of this soup of life.

The text is written as if for the non-specialist, but will be most appreciated by readers familiar with or even expert in the basics of biochemistry. There are no chemical formulas except, for the space-filling models, no equations and only a few numbers, always interesting. Important chemical insights are stated succinctly and boldly in the text. And the most elegant mechanisms are presented in an easily accessible but constantly informative way. This is a fresh and engaging book recommended to all students of biology. □

Harold P. Erickson is in the Department of Cell Biology, Duke University Medical School, Durham, North Carolina 27710, USA.

Imaging Earth

B. L. N. Kennett

Seismic Tomography. Edited by H. M. Iyer and K. Hirahara. Chapman and Hall: 1993. Pp. 842. £135, \$249.95.

METHODS of determining the three-dimensional distribution of seismic wave speeds in the Earth have come to be known as 'seismic tomography' by analogy with medical imaging techniques. But the illumination of the Earth by earthquake and man-made tremors is not nearly as uniform as the illumination of the human body by X-rays. Receivers can be placed only near the Earth's surface, so considerable ingenuity is required to get good images. Most seismic tomography today is based on the times of arrival of seismic phases, but for large-scale problems the amplitude and phase information is beginning to be used.

This compilation brings together the expertise of 58 authors in 40 articles. The articles are arranged in order of decreasing scale, starting with images of the whole mantle of the Earth and finishing with mining and engineering applications. The coverage is thorough and provides an up-to-date overview of current activity. I am unsure, however, for whom the book is intended: despite its length, it is not self-contained. Nearly every section requires an extensive knowledge of seismology for maximum benefit; so although students are likely to find the book helpful, they will need substantial help to encompass all of the articles. There is also a lot of overlap in theoretical material between many of the articles; although some parts are well integrated, the notations used in different sections are not always consistent. The editors have, however, succeeded in getting a much more extensive discussion of the reliability of results than is usually encountered.

The emphasis of most of the articles is on the detailed application of least-squares techniques for handling the arrival times of seismic waves, using iterative matrix solvers. Only a few articles look forward to suggest ways in which the quality of seismic images might be improved. As computational costs decrease, existing procedures should benefit from the inclusion of three-dimensional ray tracing. But the most dramatic improvements will probably come from careful experimental work to improve data coverage in three dimensions and from the exploitation of currently under-used information, such as the character of seismic wave forms. □

B. L. N. Kennett is in the Research School of Earth Sciences, Australian National University, Canberra, ACT 0200, Australia.

Quantum solitons in optical fibres

P. D. Drummond, R. M. Shelby, S. R. Friberg & Y. Yamamoto

The quantum-mechanical properties of long-lived optical pulses known as solitons have recently been studied theoretically and observed experimentally. These experiments suggest effective ways of reducing quantum noise in signal transmission and of performing quantum-mechanical non-demolition measurements without destroying the signal.

The first reported soliton was a water-wave in a canal, observed by John Scott-Russell¹ in 1844. A wave or 'lump' of water was seen to travel over long distances without spreading out or breaking up. Since then, numerous examples of solitons have been found in various physical, chemical and biological environments, all propagating without spreading out or breaking up. Recently, there has been great interest in solitons in optical fibres. Such solitons, predicted in 1973 (ref. 2) and first demonstrated in 1980 (ref. 3), offer advantages for high-speed, long-distance communications: propagation distances of more than one million kilometres have been demonstrated⁴. The stability of optical solitons in fibres makes them interesting not just for optical communication systems, but also for fundamental tests of quantum mechanics.

Solitons are persistent 'solitary waves' or pulses that are particularly resistant to the disruptions that ordinarily distort a pulse as it propagates. A short pulse is made up of a range of oscillation frequencies: the shorter the pulse, the broader its spectrum. In any physical medium, the propagation velocity depends on frequency, an effect known as group-velocity dispersion. By itself, this causes different portions of the pulse spectrum to travel with different velocities, leading to pulse broadening. Temporal solitons exist in nonlinear media in which the velocity also depends on the amplitude of the pulse. If the signs of these two effects are opposite a soliton can form which (in classical physics) would propagate without distortion. Spatial optical solitons form in a similar manner, opposing diffraction against self-focusing in a nonlinear medium^{5,6}.

In the case of solitons in optical fibres, the balance is between the wavelength dependence and the intensity dependence of the refractive index of the fibre waveguide. Group-velocity dispersion causes shorter wavelengths to travel faster (for the case of interest here) than longer wavelengths. Normally, the leading edge of the pulse becomes blue shifted and the trailing edge red shifted. However, the intensity dependence of the refractive index causes the phase of the pulse to be time-dependent. This results in a frequency shift towards the red on the leading edge and towards the blue on the trailing edge, tending to cancel the effect of group-velocity dispersion. When these opposing effects balance, an optical-fibre soliton is formed.

Classical arguments predict that a soliton is invariant as it propagates. However, the classical picture of solitons is not correct quantum mechanically: a true quantum soliton undergoes phase diffusion and wave-packet spreading. Quantum solitons consist of linear superpositions of solitons with different photon-number and momentum eigenstates. Each component soliton has a different phase velocity and group velocity, causing the phase and the wave-packet of quantum solitons to spread out as they propagate. This fact, implicit in an early work by Bethe in 1931 (ref. 7), has been explored recently in much more detail in the quantum theory of optical solitons⁸⁻¹¹. We now know that solitons in optical fibres are superpositions of macroscopic quantum states that correspond to clusters of 10^8 or more photons, bound together. Nevertheless, intrinsically quantum-mechanical effects emerge and can be detected even for such macroscopically large objects. Recently,

theoretical predictions of quantum phase diffusion have been confirmed experimentally¹² in one of the few experimental tests of a simple, exactly soluble quantum field theory. These experiments give direct evidence of the quantum nature of soliton propagation.

Quantum phase diffusion also provides a promising way for generating 'squeezed' states of light. Squeezed states originate in the uncertainty principle of quantum mechanics, which implies the presence of quantum noise fields even when the field is in its ground, or 'vacuum', state and no photons are present. A conventional laser beam is in a state approximately equivalent to the superposition of a classical sinusoidal electromagnetic field and these vacuum, or 'zero-point', noise fields. Such a state is known as a coherent state, and is characterized by a Poisson distribution of the number of photons, rather than by a uniquely defined photon number. The effect of this distribution is the 'shot noise' observed in the electrical current from a photodetector^{13,14}. Optical solitons, if produced by stable mode-locked lasers, are initially in coherent states of this type.

The uncertainty principle permits the reduction or 'squeezing' of photon-number noise to below the coherent state level (the 'shot-noise limit') if the uncertainty in the conjugate variable—the optical phase—is increased. Squeezed states, first predicted by Schrödinger in 1926 (ref. 15), have generated considerable recent interest^{13,16-18}. Fundamental to quantum theory, they have the practical consequence of reducing the intrinsic noise 'floor' in the detection of light. As we show below, initially coherent solitons propagating along an optical fibre evolve into squeezed states. In these, the quantum uncertainty is suppressed in one variable and enhanced in the other, allowing detection of sub-shot-noise using an interferometer.

When two optical solitons collide, the phase and position of each soliton is correlated with the photon number and momentum, respectively, of the other soliton. This has been shown theoretically in the quantum theory of soliton collisions¹⁹⁻²¹ and has now been confirmed experimentally²². These experimental results not only extend our understanding of single quantum solitons to the collision of two solitons but also provide an experimental realization of a quantum non-demolition (QND) measurement of the photon number of an optical soliton.

Measurement of an observable quantity of a quantum system is accomplished by allowing the system to interact with a measurement apparatus, or 'meter', for a period of time, and then observing the change in state of the meter. During the measurement interaction, the meter acts back to change the state of the system. Because a well designed measurement reduces the uncertainty of a measured observable, the uncertainty principle requires that the conjugate observable of the system be altered by the 'back-action' of the meter. In a QND measurement, the back-action perturbation of the measured system is strictly confined to the conjugate observable and does not affect the measured observable. Thus, an ideal QND measurement does not perturb the free evolution of the measured observable.

QND measurements were first discussed by Landau and Peierls in 1931 (ref. 23). Lately, there has been renewed interest in developing various techniques for the realization of QND

measurements^{24–27} because of their unique importance to quantum theories of measurement and practical implications for precision measurements. Optical solitons propagate and collide with unchanging pulse shape, momentum, intensity and energy. Consequently, the photon number of a soliton is not altered when its energy is measured by means of the phase shift of a soliton with which it collides. Accordingly, the conditions necessary for a QND measurement are met, and a QND measurement can be used to determine the number of photons in a soliton pulse with an accuracy better than the shot-noise limit.

Quantization of optical solitons

An optical soliton in a fibre can be described classically using the nonlinear Schrödinger equation^{2,3,5}, which is applicable to single-mode propagation in nonlinear dispersive fibres or waveguides. This is an approximate wave-equation which ignores quantum effects. In this section, we review the predictions of a quantum version of this equation, in which operators replace the classical field variables. Classically, the most unexpected behaviour is the existence of quantum solitons of discrete particle number and undetermined phase. More readily observable quantum effects include phase and position diffusion, in which classically stable observables have increasing quantum uncertainties with propagation. For soliton collisions, quantum theory predicts that information about phase and position is transferable from one soliton to another, allowing QND measurements.

The relevant quantum particles in a dielectric medium are single-particle excitations of the combined dielectric and radiation field systems, known as polaritons or dressed photons^{28,29}. We will generally use the term photon to describe such photon-like particles, even though they are distinct from the pure photons of a vacuum. For example, they travel at an average velocity $v < c$, where v is the group velocity of electromagnetic radiation in the dielectric at the carrier frequency ω_0 , and c is the speed of light in vacuum.

In a reference frame moving at the group velocity v , the corresponding quantum operator equation for the photon (that is, polariton) density amplitude $\hat{\psi}$ in dimensionless form is^{8,11,20,21,30}:

$$\frac{\partial}{\partial T} \hat{\psi}(T, X) = i \left(\pm \frac{1}{2} \frac{\partial^2}{\partial X^2} + \frac{2\hat{\psi}^\dagger \hat{\psi}}{\bar{N}} \right) \hat{\psi}(T, X) \quad (1)$$

where:

$$[\hat{\psi}(T, X), \hat{\psi}^\dagger(T, X')] = \delta(X - X') \quad (2)$$

The variable T represents elapsed time, and is therefore proportional to the distance propagated. It is defined relative to a characteristic pulse duration t_0 and group velocity dispersion $k_0'' = \partial^2 k / \partial \omega^2 /_{\omega=\omega_0}$ so that:

$$T = vt |k_0''| / t_0^2 \quad (3)$$

The variable X represents spatial location relative to the moving reference frame:

$$X = (x/v - t)/t_0 \quad (4)$$

The $(\pm)$ sign in equation (1) is positive for anomalous dispersion ($k_0'' < 0$) optical fibres that support solitons, typically at wavelengths greater than $\sim 1.4 \mu\text{m}$. The quantity $\bar{N}$ is a scaling constant inversely proportional to n_2 , the nonlinear refractive index per unit intensity:

$$\bar{N} = \frac{2|k_0''| \mathcal{A} c}{n_2 \hbar \omega_0 t_0} \quad (5)$$

Here, $\mathcal{A}$ is the effective modal cross-section of the fibre. The operator $\hat{\psi}^\dagger(T, X) \hat{\psi}(T, X)$ directly represents the photon density in dimensionless units.

The operator equations show that dressed photons propagate with a nonzero effective mass³⁰ $m = \hbar / (|k''| v^3)$ in the moving

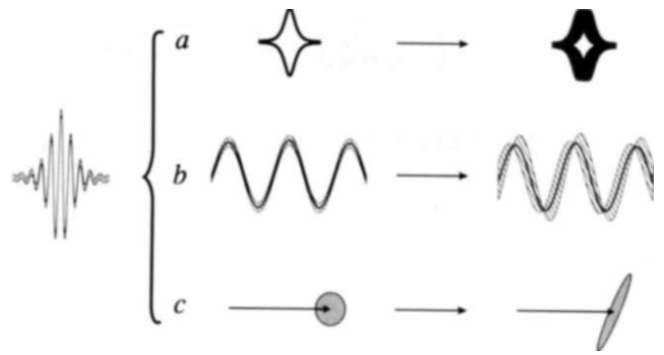


FIG. 1 Evolution of a quantum soliton. A coherent soliton, shown schematically at the left, is a superposition of a classical hyperbolic secant light pulse and quantum noise due to the zero-point fluctuations of the field. The presence of the zero-point fields imposes uncertainties in the soliton's position in time and space, its central frequency, its photon number or electric-field amplitude, and the phase of its electric field. The frequency uncertainty, coupled with the frequency dependence of the group velocity of the fibre, causes quantum spreading of the envelope of the pulse as shown in *a*. The uncertainty in photon number obeys a Poisson distribution and is responsible for the 'shot noise' observed when coherent light is detected. The photon-number uncertainty, along with the intensity dependence of the phase velocity, causes a spreading of the phase as shown in *b*. The electric field and noise are shown for an optical period. The phase and position spreading are not due simply to an increase in uncertainty, but also to correlations with the amplitude and frequency of the pulse, respectively. This results in interference effects. The destructive interference of the noise occurs at certain phases of the optical cycle, as shown in *b*. This is known as quantum noise squeezing, and can also be represented as a narrowing of the phase-space diagram of the field as shown in *c*. There, the vector represents the average electric-field amplitude and the shaded area the quantum noise. The left side is a coherent state—the superposition of the field and zero-point fluctuations with equal and uncorrelated amplitude and phase uncertainties yields a circular distribution. At the right, the soliton is shown after evolving into a squeezed state. The noise is reduced in certain directions due to the correlations, yielding an elliptical distribution with a minor axis smaller than the width of the original coherent noise distribution.

reference frame. Similar changes in effective mass are observed in electrons travelling in solids. The nonlinear term in the equation is equivalent to a localized delta-function attractive force acting between the particles. In real fibres, there are additional effects due to absorption as well as to Brillouin^{32–34} and Raman^{31,35} scattering.

The energy eigenstates of the quantum nonlinear Schrödinger equation are known from early theoretical work on interacting Bose gases^{7,36,37}. The simplest quantum soliton is an eigenstate consisting of a cluster of photons with a well-defined energy and photon-number—a photonic atom. Such a state is predicted to propagate stably, without distorting or breaking up into its component photons, just as an atom does in free space. There are forces predicted to act between two such quantum solitons, as there are between two atoms. Quantum states of this type can be pulses of macroscopic size, with 10^9 or more photons. Despite this, an isolated quantum soliton-number eigenstate has yet to be experimentally observed.

Fibre soliton experiments^{12,22} start with the input field in a coherent state^{38,39}, rather than in a state of well-defined particle number. In a coherent state, the quantum fluctuations of the electric field are identical to those of a vacuum state. As shown in Fig. 1, the variance is the same for either amplitude or phase fluctuations. Thus, the photon number N must have fluctuations such that $\langle (\Delta N)^2 \rangle = \langle N \rangle$. This means that an input field will be a superposition of quantum solitons of different photon number, along with some unbound 'continuum' photons.

Even for superpositions of this type, there are large differences between classical and quantum theories. We first note that the classical equations predict^{2,3,5} that a coherent pulse will propagate essentially unchanged if injected with an initial amplitude of $\sqrt{N}/2 \operatorname{sech}(X)$. The time-dependent classical solution is given by:

$$\psi_0(T, X) = \sqrt{N}/2 \operatorname{sech}(X) \exp[i(\theta_0 + T/2)] \quad (6)$$

where θ_0 is the initial phase. The parameter $\bar{N}$ corresponds to the mean photon number $\langle N \rangle$. Thus, the only change predicted classically is that the total phase $\theta(T) = \theta_0 + T/2$ increases linearly with distance down the fibre.

The classical result of a soliton as a phase-shifted, undistorted pulse never occurs in quantum theory. Even for a quantum soliton with a well defined photon number, the centre-of-mass position obeys the Heisenberg uncertainty principle. Position and momentum are conjugate observables and thus not simultaneously measurable, so a soliton with some given velocity cannot have a well-defined position at all times. This holds as well for the conjugate observables of phase and photon number. Momentum and photon number are invariant during propagation. Their conjugates (position and phase) are predicted to show a quantum-mechanical time dependence—leading to increasing fluctuations as time progresses. This is the quantum diffusion of position and phase.

In order to understand quantum phase diffusion better, consider a simplified model of the input coherent pulse as a superposition of quantum solitons of different photon number. The classical phase shift of a soliton with N photons at elapsed time T is:

$$\theta_N(T) = \frac{TN^2}{2\bar{N}^2} = \frac{x|k_0'|N^2}{2t_0\bar{N}^2} \quad (7)$$

This phase shift can also be calculated directly from quantum theory, using the fact that the optically measured phase⁴⁰ is the difference between the quantum phase of the N -photon and the $(N+1)$ -photon states. As many photon number states are superimposed to give an initially coherent soliton, this leads to a corresponding diffusion of phase shifts after an elapsed time T . The enhanced phase fluctuations can be roughly estimated from the known fluctuations in photon number N , giving $\langle \Delta\theta(T)^2 \rangle \approx T^2/\bar{N}$. More rigorously, the initial coherent-state phase fluctuations should also be included. This gives the final result that^{8,11,41}:

$$\langle \Delta\theta(T)^2 \rangle = (0.607 + T^2)/\bar{N} \quad (8)$$

A similar result holds for the quantum diffusion of the centre-of-mass position.

In real soliton experiments, a number of additional factors must be taken into account. Experimental local oscillator measurements⁴² do not correspond exactly to an idealized soliton phase measurement^{10,11}, and the initial pulse-shape is seldom precisely an ideal sech pulse. The phase also experiences thermal disturbances due to Brillouin and Raman scattering. Even so, the numerically predicted and experimentally observed phase fluctuations are rather close to those of equation (7).

As a result of phase diffusion, the distribution of the amplitude fluctuations narrows, or squeezes, for certain phases. This leads to quantum noise reduction, as illustrated in Fig. 1. Its physical origin is the correlation of photon-number and phase for each component of the superposition. Quantum noise reduction due to soliton squeezing has possible applications in low-noise measurements.

We now turn to soliton collision experiments, which have been analysed using quantum inverse scattering methods^{19,21}. In fact, it is possible to understand soliton collisions using classical inverse scattering results⁵ together with stochastic noise sources representing vacuum fluctuations³¹. After two solitons collide, the position and phase of one soliton is modified by an amount that depends on the velocity and photon number of the other

soliton. Quantitatively, we find:

$$\Delta\theta_p^{\text{COL}} = 2 \left(\tan^{-1} \left[\frac{N_p + N_s}{\bar{N} |V_p - V_s|} \right] - \tan^{-1} \left[\frac{N_p - N_s}{\bar{N} |V_p - V_s|} \right] \right) \quad (9)$$

For large relative velocity differences, the result for the phase reduces to:

$$\Delta\theta_p^{\text{COL}} = \frac{4N_s}{\bar{N} |V_p - V_s|} \quad (10)$$

Where $\Delta\theta_p^{\text{COL}}$ is the phase shift experienced by a probe soliton, V_p and V_s are the velocities of the probe- and signal-soliton in dimensionless units, and N_p and N_s are the corresponding numbers of photons. The physical explanation of this effect is that as the solitons pass through each other, they experience a more intense field. Due to the nonlinearity of the refractive index, this increases the phase-shift experienced by each soliton. The phase-shift is proportional to the photon number of the other soliton—but is smaller at large relative velocities, as this reduces the interaction time.

A probe soliton phase measurement after a collision of probe and signal solitons allows the photon number of the signal soliton to be measured, as $\Delta\theta_p^{\text{COL}}$ is proportional to N_s . The significant feature of this interaction is that, neglecting Raman

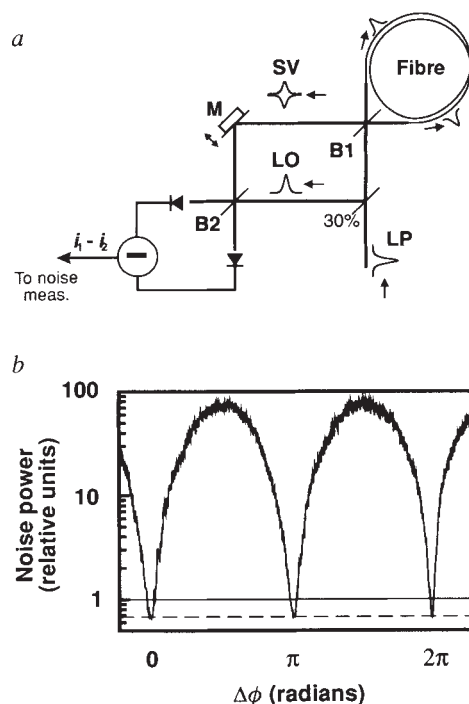


FIG. 2 a, Schematic diagram of the experiment to observe squeezed solitons. A laser pulse (LP) is split and launched into the two ends of a fibre as solitons. Propagation squeezes both solitons, and when they are recombined at beamsplitter B1, they produce a pulse of pure squeezed noise—the squeezed vacuum (SV). The classical portions of the soliton combine at the other port of the beamsplitter to produce an intense pulse which can be used as a local oscillator (LO) for detection of the squeezing. The squeezing is observed by combining the intense pulse (with a phase shift determined by the position of the translating mirror M) with the squeezed noise at the beamsplitter B2. The resulting pulses are detected by two photodetectors and the difference of their photocurrents exhibits the phase-sensitive squeezed noise. The result of such a noise measurement is displayed in b as a function of the phase shift of the intense pulse, $\Delta\phi$. The vertical axis is the photocurrent noise power plotted relative to the coherent state limit. The dashed line is at a noise level of 0.66, indicating that the photocurrent noise is suppressed to 34% below the classical minimum. This, plus the observed phase dependence of the noise, provides clear evidence of squeezing.

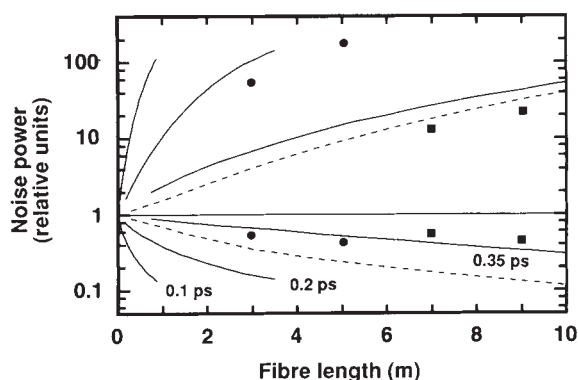


FIG. 3 Experimental soliton-squeezing data (corrected for the experimental detection efficiency) plotted for pulses of 0.35 ps (squares) and 0.2 ps (circles). Both the noise minima and maxima for pulse lengths of 0.35, 0.2 and 0.1 ps are compared with theoretical predictions (solid lines) that take into account the thermal noise in a fibre at liquid-nitrogen temperatures. Agreement between experiment and theory is better for longer pulses for reasons discussed in the text. For shorter pulses, thermal noise effects are weaker, and better quantum noise suppression is predicted for shorter propagation distances. Hypothetical squeezing performance for 0.35-ps pulses in the absence of thermal noise is shown by a dashed line.

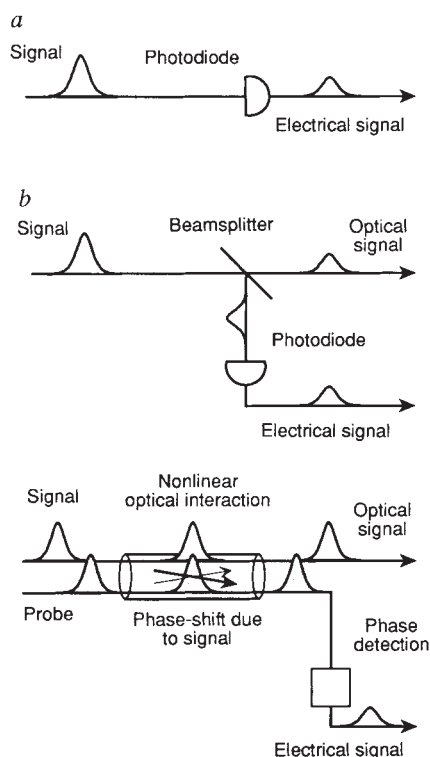


FIG. 4 Optical signal measurement techniques and QND. The usual method to measure an optical signal, illustrated in *a*, is to completely absorb it in a photodiode to produce an electrical current. A more sophisticated approach, shown in *b*, is to only partially absorb the optical signal. A quantum non-demolition (QND) measurement, performed as shown in *c*, completely preserves the signal and adds no noise. For a soliton QND measurement, a signal and a probe soliton interact by means of a Kerr nonlinearity in a single-mode optical fibre. The interaction causes a phase shift of the probe proportional to the energy (photon number) of the signal. There is no transfer of energy, so the photon number is not altered. The phase shift of the probe soliton (which is absorbed) can then be determined, measuring, but not perturbing, the photon number of the signal soliton.

interactions, neither the photon number nor the velocity of either soliton is changed during the collision. In contrast, all other photon-number measurement techniques used in practice modify the photon number. Thus, soliton collisions allow a back-action evading or quantum non-demolition measurement of the soliton photon number.

Observation of squeezed optical solitons

Squeezed states of light, as described above, have intrinsic quantum noise fluctuations which vary in magnitude during an optical cycle. At certain phases of the cycle, the fluctuations are smaller than those of a coherent state, or equivalently, the shot-noise limit. The transformation of an initially coherent soliton into a squeezed soliton was recently observed experimentally¹². Laser pulses of a fraction of picosecond (10^{-12} seconds, ps) duration and with centre wavelengths in the negative group-velocity dispersion region of a single-mode optical fibre were produced by a mode-locked colour centre laser. These pulses were, to good approximation, in coherent states. Squeezing of these pulses was studied for fibre lengths of 3–10 m, corresponding to normalized propagation distances of $T=2-8$.

As an initially coherent optical soliton propagates in a fibre, correlations build up between its intensity and phase. Squeezing occurs due to destructive interference of the noise at certain phases of the optical cycle, as shown in Fig. 1*b*. Because intensity-phase correlations do not change the intensity, they do not directly reduce the intensity noise (that is, the photon-number fluctuations). Detectors which respond to intensity, but not to phase, do not reveal this type of squeezing. Therefore, in experiments, squeezing is transformed into a reduction of the photon-number fluctuations by interference.

The required transformation is a phase shift of the squeezed noise with respect to the average optical field. This can be viewed either as sliding the noise minimum to the peak of the optical sine wave (Fig. 1*b*), or as an appropriate rotation of the noise ellipse (Fig. 1*c*). To accomplish this, the squeezed quantum noise was 'dissected' from the soliton⁴³, as outlined in Fig. 2*a*. Two solitons were launched into opposite ends of a fibre. After leaving the fibre, the two solitons, now squeezed, interfered at a beamsplitter. The classical fields of the two solitons are derived from a single laser pulse, so they both have an identical phase and therefore exhibit complete destructive interference at one of the outputs of the beamsplitter. The quantum fluctuations acquired by the two solitons, however, are completely independent, so they combine at that output. This yields a pure squeezed-noise-field pulse known as a 'squeezed vacuum'—a noise burst with no average electric field but with the correlations and noise interferences induced by propagation through the fibre. In the experiment, this squeezed noise pulse was recombined with the intense pulse resulting from constructive interference of the two solitons at the other beamsplitter output. With the correct phase shift, this yielded a pulse with reduced photon-number noise. By varying the phase, the full phase dependence of the squeezed noise was observed.

The measured squeezed soliton noise is shown in Fig. 2*b*, with the mean-square noise of the squeezed light plotted as a function of the phase difference between the squeezed vacuum and the intense pulse. The noise is plotted relative to the shot-noise level—values less than one are in the nonclassical regime. Phase-dependent nonclassical noise such as this is the unmistakable signature of a squeezed state, and is the first quantum effect observed for solitons. The photocurrent noise was 34% below the shot-noise level for the case of a 0.2 ps soliton pulse and 5 m of fibre. Taking into account the overall detection efficiency of $\eta \sim 0.65$, it can be inferred that the mean-square optical field fluctuations were reduced below the shot-noise level by 52%.

A fundamental limitation to quantum noise reduction is the phase noise imposed by thermal refractive-index fluctuations in optical fibres. The most significant for fibre squeezing^{34,44,45} is guided acoustic wave Brillouin scattering, or GAWBS^{32,33}. It

is caused by ultrasonic waves confined by the cylindrical fibre structure, and is in itself quite interesting. For very short pulses, Raman scattering can also contribute noise.

To describe soliton squeezing fully, the quantum theory of soliton propagation must be modified to include thermal noise sources³⁵. Predicted results including such modifications and the experimentally observed results are shown in Fig. 3 as a function of the length of the fibre. Quantum noise suppression continuously improves as the pulse is shortened. Notice that GAWBS noise reaches the shot-noise level after a soliton propagates about 10 m, so the solitons must be short enough and have enough peak power to ensure that squeezing takes place in less distance. For typical fibres, sub-picosecond pulses are required. If we consider 0.1-ps pulses, for example, fibre lengths can be made so short that thermal noise never significantly degrades the squeezing, as is evident from Fig. 3.

The shorter pulses in the experiments did not produce the expected noise reduction, probably because of imprecise matching of the two solitons excited in the fibre. If the two pulses differ slightly, the two contributions to the squeezed noise pulse produced at beamsplitter B1 do not line up exactly in phase, and the noise minimum is partially obscured. With longer pulses, matching was more easily achieved and the agreement with theory is better.

For sub-picosecond pulses, soliton squeezing may be a convenient source of squeezed light. In some applications, however, it may be more convenient to measure precisely the photon-number noise with quantum non-demolition measurement, rather than to directly reduce it with squeezing.

Quantum non-demolition measurements

Knowledge about the world comes through observation and measurement, making the nature of measurement of fundamental importance. The sometimes startling differences between classical and quantum theories of measurement thus prove very intriguing. In classical physics a measurement is not expected to change the properties of a system being measured, whereas in quantum physics it is.

Everyday experience gives mixed results, suggesting that measurement changes some quantities but not others. Optical signals, for example, are usually measured by photodetectors such as the eye or its electrical equivalent as in Fig. 4a. This produces neural or electrical signals while simultaneously demolishing the optical signal. A more sophisticated optical measurement, as shown in Fig. 4b, may absorb only a small part of the optical signal but still alters it. Our general experience, however, is that of classical physics. We do not expect the flight of a tennis ball, for example, to be affected by spectators.

It is perhaps surprising that quantum solitons—optical pulses with well defined quantum-mechanical properties—can be measured in a way very similar to a classical measurement. As described previously, this can be done by colliding two solitons with different frequencies. This causes the phase of, say, a probe soliton to become correlated with the photon number of a signal soliton. A measurement of the phase shift of the probe then gives the photon number of the signal, as shown in Fig. 4c. After being measured, the signal soliton propagates freely, its energy unchanged. Heisenberg's uncertainty relationship is not violated—the measurement causes an increased uncertainty only in the phase of the signal soliton.

For both the photon-number and the momentum eigenstates of an optical soliton, it is possible to perform two or more back-action evading measurements that yield identical outcomes and do not change the quantum-mechanical state of the soliton. These are known as quantum non-demolition (QND) measurements^{23–27}. Several optical QND measurements have been proposed theoretically^{46–48} and partial QND measurements have been demonstrated experimentally by several groups^{49–51}. However, two successive measurements on the same quantum state have yet to be achieved.

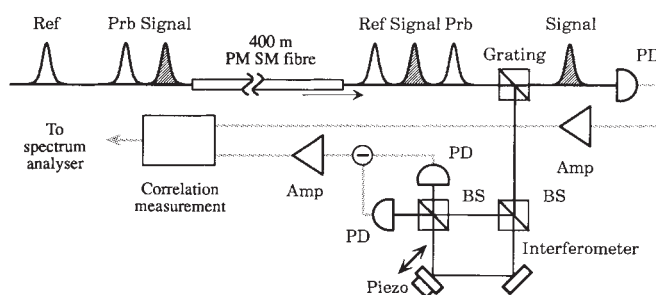


FIG. 5 A schematic diagram of the quantum non-demolition (QND) measurement. A signal soliton (Sig) at 1,460.7 nm and identical probe (Prb) and reference (Ref) solitons at 1,455 nm propagate in a 400-m polarization-maintaining (PM), single-mode (SM) fibre having anomalous group-velocity dispersion. The signal and probe solitons collide in the fibre, causing the phase of the probe soliton to shift in response to the photon number of the signal soliton. After leaving the fibre, the solitons are separated by a dispersing grating. A photodiode (PD) records the photon-number fluctuations of the signal soliton and an interferometer (with one arm slightly longer than the other so that the probe and reference pulses overlap) measures the phase-shift fluctuations of the probe soliton. Correlation measurements with a radio-frequency spectrum analyser show the fluctuations to be correlated, indicating a partial QND measurement.

The unique characteristics of quantum solitons promise to make successive QND measurements possible. Except for solitons, all optical pulses experience broadening due to group-velocity dispersion in a dielectric medium. Thus, the photon number in any given time interval (the intensity of the pulse) is not a well defined quantity. Much the same happens for continuous-wave optical signals, unless they are strictly monochromatic. However, quantum solitons can propagate with a well defined photon-number state, so two sequential measurements on the same quantum state can easily be carried out. As subsequent measurements can take place in a single optical fibre, losses can nearly be eliminated.

A partial soliton QND measurement of photon number was recently demonstrated experimentally²². Three jitter-free optical pulses—a signal, a probe and an identical reference—were extracted from single colour-centre-laser pulses spectrally broadened to a 10-nm bandwidth by self-phase modulation in an optical fibre. These were launched into a 400-m negative GVD, single-mode, polarization-preserving fibre, yielding a signal soliton at 1,460.7 nm and identical probe and reference solitons at 1,456.0 nm, all with ~1.0-nm bandwidths and 2.2-ps pulse widths. Timing was adjusted so the signal and probe collided halfway along the fibre. The reference soliton, offset 30 ps from the probe, propagated without colliding. The three closely-spaced solitons moved more quickly through the fibre than it could change due to thermal noise processes. This eliminated most of the detrimental effects of GAWBS noise which usually plagues quantum optical measurements in optical fibres^{32,33,44}.

The solitons, after colliding and leaving the fibre, were separated with a diffraction grating. The signal soliton, directed into a photodetector, gave a current proportional to its photon number. The probe and reference solitons, interfering in a soliton-collision interferometer^{52,53} as shown in Fig. 5, gave a current corresponding to the phase of the probe relative to that of the reference.

The results were as predicted by equations (9) and (10); the phase shift of the probe was proportional to the photon number of the signal soliton. For example, when the signal soliton was switched on and off, the probe solitons's phase was modulated by 1.2 radians (70°). In the quantum regime, the fluctuations of the photon number of the signal were observed to be correlated with the phase fluctuations of the probe. Furthermore, the signal

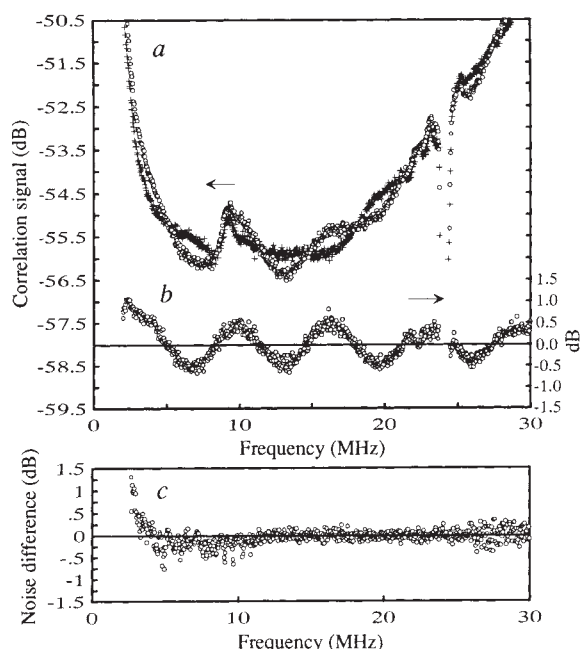


FIG. 6 Experimental results verifying the QND nature of the soliton collision measurement. The correlation between the shot noise of the signal soliton and the phase noise of the probe soliton is indicated by the modulation of the two radio-frequency spectrum-analyser traces in *a*. Subtracting the two for clarity gives *b*. The lack of excess noise due to the measurement is shown in *c*, where the data points are the difference of the photon-number fluctuations of the signal soliton and a quantum-limited calibration signal. Above 12 MHz, they cluster about the standard quantum limit (solid line). Taken together, these results show the photon number of the signal soliton and the phase of the probe soliton to be quantum correlated.

soliton's photon-number fluctuations were unchanged by the collision and shot-noise limited, showing that no excess noise was introduced back into the photon number of the signal.

The results of the experiment are shown in Fig. 6. Noise correlations were found by adding and then subtracting the two noise signals—differing results indicated a correlation. In the experiment, a delay line caused the signal soliton shot-noise fluctuations to alternate in sign depending on the frequency. When the two noise signals were summed, correlations caused a modulation of the noise spectrum. Such a modulation, observed with a radio-frequency spectrum analyser, is seen in the two sets of experimental data shown in Fig. 6*a*. Figure 6*b* shows the modulation more clearly, after the data sets are subtracted to eliminate background variations. The results of the check for excess noise in the signal soliton, carried out immediately after the correlation measurements, are shown in Fig. 6*c*. Above 12 MHz, the average difference between the noise of the signal solitons and a quantum-limited calibration signal was less than 0.1 dB, indicating no excess noise due to either the collision or external factors.

The experiment showed that soliton-collision phase measurements satisfy the criterion for a QND measurement. The QND read-out (the phase shift of the probe soliton) and the QND observable (the photon number of the signal soliton) were correlated; the measurement was back-action evading (no noise was introduced into the signal soliton's photon number) and the QND observable was not changed either by measurement, by propagation, or by external noise. Thus, for a soliton initially in a coherent state, a measurement as described here puts the soliton in a near photon-number state—a partial QND measurement causes photon-number squeezing. An identical second measurement on the same signal soliton, if correlated with the first measurement, could verify that both measurements gave the same result. In principle, this would be a full QND measurement that would leave the quantum-mechanical state of the soliton unaltered. It would be a classical-like measurement in that it would determine the soliton's photon number without changing it.

Outlook

We have described recently demonstrated quantum soliton effects—squeezing and QND measurements—in an optical fibre. In both cases, the conjugate observables of photon number and phase were studied. However, squeezing and QND measure-

ments for the other pair of conjugate observables, position and momentum, are yet to be observed.

Solitons can be transmitted over long distances with relatively little distortion, and have the ability to transmit information in several different wavelength-channels with minimum interaction. Consequently, they have become important candidates for long-distance terrestrial and trans-oceanic telecommunications^{54,55}. Both quantum and classical spreading of soliton pulse positions are sources of error in soliton communication systems. With present technology, timing jitter caused by amplifier noise⁵⁶ (from the amplifiers used to overcome the small fibre loss) is likely to be more important over long distances than quantum diffusion of the pulse position. Exploiting correlations between position and momentum is one possible approach to reducing such errors.

Because solitons interact in a particle-like manner, preserving their identity when perturbed, they are also likely to be important for viable optical switching devices or logic gates^{57,58}. Non-soliton implementations often suffer from pulse distortions which cause incomplete switching. From this perspective, a soliton collision QND experiment is similar to an optical switching device—information coded on the signal soliton by means of its presence or by its photon number is transferred to the probe soliton^{59,60}. To a good approximation this realizes a reversible quantum logic gate^{61–63}, a device of considerable theoretical interest. Quantum effects in the propagation and collisions of solitons may ultimately determine the limits to performance of optical switching devices. Such limits are only now beginning to be defined⁶⁴.

It is interesting to consider other physical systems that have optical solitons with stronger nonlinearities. In a mode-locked dye laser or solid-state laser, for example, the coexistence of a Kerr nonlinearity and a negative group velocity dispersion inside the laser cavity produces soliton-like pulses⁶⁵. Such mode-locked laser pulses inside a cavity have much higher peak intensities than optical solitons in a fibre, so more efficient squeezing and QND measurements might be realized⁶⁶.

In most transparent media, the nonlinear and linear dispersive effects have the same sign, and solitons do not form. It has been suggested that pulses known as gap solitons may nevertheless be formed in such media by exploiting the dispersion of a spatially periodic refractive-index modulation fabricated into the nonlinear medium⁶⁷. These solitons have yet to be observed, but

due to the large nonlinear phase shifts predicted for very small propagation distances, they might be very useful for quantum noise squeezing or for optical switching applications.

Self-induced transparency in resonant two-level systems produced a slightly different type of optical soliton, mathematically described by a Maxwell–Bloch equation⁶⁸. The leading edge of the pulse excites ground-state atoms to the excited state, then the trailing edge of the pulse stimulates transitions back to the ground state. In this way, the optical pulses propagate without being attenuated or distorted in the medium. As self-induced transparency solitons exist in resonant two-level systems, large nonlinearities are expected, even for solitons with low intensities^{20,21}. Resonant two-level systems also are expected to support spatial solitons with quantum properties similar to those described here, but with larger nonlinearities⁶⁹. Quantum solitons of a different variety—topological solitons—are predicted in a number of model quantum field theories used in solid-state and particle physics⁷⁰. It is possible that our methods could be used to directly investigate these solitons, whose quantum properties are as yet untested by experiment.

Coherent optical communications and erasable optical data storage are increasingly important technologies. As these progress to the regime of very high bandwidths and data densities, vacuum fluctuations will become a limitation on the signal-to-noise ratios that can be achieved. A practical means for generation of squeezed light pulses would overcome this limitation^{71–73}. The use of short pulses, as described here, would overcome the thermal-noise problems that so far have limited the applicability of fibres for squeezing and QND. Technological benefits of quantum noise suppression depend upon the trade-off between cost, complexity and improvements in measurement sensitivity. In this respect, solitons in optical fibres offer a simple and robust technique for nonclassical light generation and application. □

P. D. Drummond is at the Department of Physics, University of Queensland, Queensland, Australia 4067. R. M. Shelby is at the IBM Almaden Research Center, 650 Harry Rd, San Jose, California 95120, USA. S. R. Friberg and Y. Yamamoto are at the NTT Basic Research Laboratories, Musashino-shi, Tokyo 180, Japan. Y. Yamamoto is also at the Edward L. Ginzton Laboratory, Stanford University, Stanford, California 94305, USA.

1. Scott-Russell, J. *Proc. R. Soc. Edinb.* **1**, 433–434 (1844).
2. Hasegawa, A. & Tappert, F. *Appl. Phys. Lett.* **23**, 142–144 (1973).
3. Mollenauer, L. F., Stolen, R. H. & Gordon, J. P. *Phys. Rev. Lett.* **45**, 1095–1098 (1980).
4. Nakazawa, M., Yamada, E., Kubota, H. & Suzuki, K. *Electron. Lett.* **27**, 1270–1272 (1991).
5. Zakharov, V. & Shabat, A. *Zh. eksp. teor. Fiz.* **61**, 118–134 (1971); also published as *Soviet Phys. JETP* **34**, 62–69 (1972).
6. Barthelemy, A., Maneuf, S. & Froehly, C. *Opt. Commun.* **55**, 201–206 (1985).
7. Bethe, H. A. *Z. Phys.* **71**, 205–226 (1931).
8. Carter, S. J., Drummond, P. D., Reid, M. D. & Shelby, R. M. *Phys. Rev. Lett.* **58**, 1841–1843 (1987).
9. Drummond, P. D. & Carter, S. J. *J. opt. Soc. Am.* **B4**, 1565–1573 (1987).
10. Lai, Y. & Haus, H. A. *Phys. Rev. A* **40**, 844–853 (1989).
11. Haus, H. A. & Lai, Y. *J. opt. Soc. Am.* **B7**, 386–392 (1990).
12. Rosenbluh, M. & Shelby, R. M. *Phys. Rev. Lett.* **66**, 153–156 (1991).
13. Walls, D. F. *Nature* **306**, 141–146 (1983).
14. Shapiro, J. H. *IEEE J. Quant. Electron.* **QE-21**, 237–250 (1985).
15. Schrödinger, E. *Naturwissenschaften* **14**, 664 (1926).
16. *J. opt. Soc. Am.* **B4**, (Special issue on squeezed states of the electromagnetic field: eds Kimble, H. J. & Walls, D. F.) 1450–1741 (1987).
17. *J. mod. Optics* **34**, (Special issue on squeezed states light: eds Loudon, R. & Knight, P. L.) 709–1020 (1987).
18. *Appl. Phys.* **B55**, (Special issue on squeezed states and nonclassical light: eds Giacobino, E. & Fabre, C.) 189–303 (1992).
19. Haus, H. A., Watanabe, K. & Yamamoto, Y. *J. opt. Soc. Am.* **B6**, 1138–1148 (1989).
20. Watanabe, K., Nakano, H., Honold, A. & Yamamoto, Y. *Phys. Rev. Lett.* **62**, 2257–2260 (1989).
21. Watanabe, K. & Yamamoto, Y. *Phys. Rev. A* **42**, 1699–1702 (1990).
22. Friberg, S. R., Machida, S. & Yamamoto, Y. *Phys. Rev. Lett.* **69**, 3165–3168 (1992).
23. Landau, L. & Peierls, R. *Z. Phys.* **69**, 56–59 (1931).
24. Braginsky, V. B. & Vorontsov, Y. I. *Usp. Fiz. Nauk* **114**, 41–53 (1974); also published as *Sov. Phys. -Usp.* **17**, 644–650 (1975).
25. Caves, C. M., Thorne, K. S., Drever, R. W. P., Sandberg, V. D. & Zimmermann, M. *Rev. mod. Phys.* **52**, 341–392 (1980).
26. Braginsky, V. B., Vorontsov, Yu. I. & Thorne, K. S. *Science* **209**, 547–557 (1980).
27. Braginsky, V. B. & Khalili, F. Ya. *Quantum Measurement* (Cambridge Univ. Press, 1992).
28. Hopfield, J. J. *Phys. Rev.* **112**, 1555–1567 (1958).
29. Drummond, P. D. *Phys. Rev. A* **42**, 6845–6857 (1990).
30. Yurke, B. & Potasek, M. J. *J. opt. Soc. Am.* **B6**, 1227–1238 (1989).
31. Drummond, P. D. & Hardman, A. D. *Europhys. Lett.* **21**, 279–284 (1993).
32. Shelby, R. M., Levenson, M. D. & Bayer, P. W. *Phys. Rev. Lett.* **54**, 939–942 (1985).
33. Shelby, R. M., Levenson, M. D. & Bayer, P. W. *Phys. Rev. B* **31**, 5244–5252 (1985).
34. Shelby, R. M., Drummond, P. D. & Carter, S. J. *Phys. Rev. A* **42**, 2966–2976 (1990).
35. Carter, S. J. & Drummond, P. D. *Phys. Rev. Lett.* **67**, 3757–3760 (1991).
36. Lieb, E. H. & Liniger, W. *Phys. Rev.* **130**, 1605–1616 (1936).
37. Yang, C. N. *Phys. Rev.* **168**, 1920–1923 (1967).
38. Glauber, R. J. *Phys. Rev.* **131**, 2766–2788 (1963).
39. Sudarshan, E. C. G. *Phys. Rev. Lett.* **10**, 277–279 (1963).
40. Pegg, D. T. & Barnett, S. M. *Phys. Rev. A* **39**, 1665–1675 (1989).
41. Kaup, D. J. *Phys. Rev. A* **42**, 5689–5694 (1990).
42. Drummond, P. D., Carter, S. J. & Shelby, R. M. *Opt. Lett.* **14**, 373–375 (1989).
43. Shirasaki, M. & Haus, H. A. *J. opt. Soc. Am.* **B7**, 30–34 (1990).
44. Milburn, G. J. *J. opt. Soc. Am.* **B4**, 1476–1489 (1987).
45. Levenson, M. D. & Shelby, R. M. *J. Mod. Opt.* **34**, 775–792 (1987).
46. Milburn, G. J. & Walls, D. F. *Phys. Rev. A* **28**, 2065–2070 (1983).
47. Yurke, B. *J. opt. Soc. Am.* **B5**, 732–738 (1985).
48. Imoto, N., Haus, H. A. & Yamamoto, Y. *Phys. Rev. A* **32**, 2287–2292 (1985).
49. Levenson, M. D., Shelby, R. M., Reid, M. & Walls, D. F. *Phys. Rev. Lett.* **57**, 2473–2476 (1986).
50. La Porta, A., Slusher, R. E. & Yurke, B. *Phys. Rev. Lett.* **62**, 28–31 (1989).
51. Grangier, P., Roch, J.-F. & Roger, G. *Phys. Rev. Lett.* **66**, 1418–1421 (1991).
52. Friberg, S. R. in *Coherence and Quantum Optics VI* (eds Eberly, J. H., Mandel, L. & Wolf, E.) 327–331, (Plenum, New York, 1990).
53. Sakai, Y., Hawkins, R. J. & Friberg, S. R. *Opt. Lett.* **15**, 239–241 (1990).
54. Taylor, J. R. *Optical Solitons: Theory and Experiment* (Cambridge Univ. Press, 1992).
55. Haus, H. A. *Proc. IEEE* **81**, 970–983 (1993).
56. Gordon, J. P. & Haus, H. A. *Opt. Lett.* **11**, 665–667 (1986).
57. Islam, M. N. *Ultrafast Fiber Switching Devices and Systems* (Cambridge Univ. Press 1992).
58. *Opt. Quant. Electron.* **24**, (Special issue on all-optical switching using solitons: ed. Wright, E.) S1215–S1336 (1992).
59. Friberg, S. R. & Jiang, W. B. in *Photonic Switching II*, (eds Tada, K. & Hinton, H. S.) 138–141 (Springer, Berlin, 1990).
60. Friberg, S. R. *Appl. Phys. Lett.* **63**, 429–431 (1993).
61. Bennett, C. H. *Scient. Am.* **257**(5), 88–96 (1987).
62. Yamamoto, Y., Kitagawa, M. & Igeta, K. in *3rd Asia-Pacific Physics Conference* (eds Chan, Y. W., Leung, A. F., Yang, C. N. & Young, K. 779–799 (World Scientific, Singapore, 1989).
63. Milburn, G. J. *Phys. Rev. Lett.* **62**, 2124–2127 (1987).
64. Sanders, B. C. & Milburn, G. J. *Phys. Rev. A* **45**, 1919–1923 (1992).
65. Avramopoulos, H., French, P. M. W., Williams, J. A. R., New, C. H. & Taylor, J. R. *IEEE J. Quant. Electron.* **QE-24**, 1884–1892 (1988).
66. Watanabe, K., Ishida, Y., Yamamoto, Y., Haus, H. A. & Lai, Y. *Phys. Rev. A* **42**, 5667–5674 (1990).
67. Martijn de Sterke, C. & Sipe, J. E. *Opt. Lett.* **14**, 871–873 (1989).
68. McCall, S. & Hahn, E. L. *Phys. Rev.* **183**, 457–485 (1969).
69. Chiao, R. Y., Deutsch, I. H., Garrison, J. C. & Wright, E. W. in *Frontiers in Nonlinear Optics* (eds Walther, H., Korotkev, N. & Scully, M. O. 151–182 (IOP, Philadelphia, 1993).
70. Coleman, S. *Aspects of Symmetry* (Cambridge Univ. Press, 1985).
71. Slusher, R. E. & Yurke, B. *IEEE J. Lightwave Technol.* **8**, 466–477 (1990).
72. Yamamoto, Y. et al. in *Progress in Optics XXVIII* (ed. Wolf, E.) 89–179 (North-Holland, Amsterdam, 1990).
73. Drummond, P. D. & Caves, C. M. in *Quantum Measurements in Optics* (eds Tombesi, P. & Walls, D. F.) 279–294 (Plenum, New York, 1992).

ACKNOWLEDGEMENTS. P. D. D. acknowledges the contributions of S. J. Carter and A. D. Hardman to the calculations of squeezing in solitons. R. M. S. acknowledges the contributions of M. Rosenbluh and E. A. Whitaker to the measurements of soliton squeezing. S.R.F. and Y.Y. acknowledge the contributions of S. Machida and K. Watanabe to the soliton QND measurements. Research at IBM was supported in part by the United States Office of Naval Research.

Adrenergic signals direct rhythmic expression of transcriptional repressor CREM in the pineal gland

Jörg H. Stehle^{*}, Nicholas S. Foulkes[†], Carlos A. Molina[†],
Valérie Simonneaux^{*}, Paul Pévet^{*} & Paolo Sassone-Corsi^{†‡}

^{*} Neurobiologie des Fonctions Rythmiques et Saisonnieres, CNRS URA 1332, Université Louis Pasteur, 12 rue de l'Université, 67000 Strasbourg, France

[†] Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, U184 de l'INSERM, Faculté de Médecine, 11 rue Humann, 67085 Strasbourg, France

Transcription factor CREM appears to play a key physiological and developmental role within the hypothalamic–pituitary–gonadal axis. This axis is modulated by the pineal hormone melatonin, whose production is in turn driven by the endogenous clock. There is striking circadian fluctuation of a novel CREM isoform, ICER, which is expressed at high levels during the night. ICER is generated from an alternative, intronic promoter and functions as a powerful repressor of cyclic AMP-induced transcription. Rhythmic adrenergic signals originated by the clock direct ICER expression by stimulation of the cAMP signal transduction pathway.

SEASONAL and circadian rhythms are central to most biological systems^{1–3}. In vertebrates, endogenous clock mechanisms regulate many aspects of physiology and metabolism^{1,4}. Crucial elements for the synchronization of biological rhythms in mammals are the pineal gland^{5,6} and the suprachiasmatic nucleus (SCN)^{7,8}. Environmental lighting conditions are transduced by the pineal gland from a neuronal to an endocrine message, the rhythmic secretion of melatonin^{5,6,9}. This hormone synthesis is controlled by the SCN, being elevated at night and low during the day^{6,9}. The cAMP-dependent signal transduction pathway serves as a relay to stimulate melatonin synthesis^{9–11}. Thus, from neuronal pathways, including the retina and the SCN, the pineal gland acts as a temporal regulator for the function of the hypothalamic–pituitary–gonadal axis^{6,12–14}.

Nuclear targets of the cAMP-dependent pathway are transcription factors that bind to promoter elements termed CREs (cAMP-responsive elements)^{15–17}. Among these factors, the products of the CREM gene¹⁸ have been implicated in the regulation of neuroendocrine processes^{19–21}. CREM encodes a family of both activators and repressors of cAMP-induced transcription by alternative splicing^{19,20,22}.

Here we document a striking circadian fluctuation of CREM expression in the pineal gland. During the night, a novel CREM isoform, termed ICER, is expressed at high levels. This isoform is generated from an alternative, intronic promoter and functions as a repressor of cAMP-induced transcription. Its expression is regulated by the SCN through adrenergic stimulation of the cAMP-dependent signalling system.

Day–night switch in CREM expression

We investigated circadian CREM expression in the pineal gland. Mid-sagittal brain sections from rats killed at consecutive time points throughout the light–dark cycle were hybridized with a CREM-specific probe (Fig. 1a, probe 1). A large difference in CREM expression is observed between day and night; an intense signal is visible in the pineal gland when animals are killed during the middle of the dark period (Fig. 2a). Similar data were obtained with golden hamsters (not shown). In contrast, no specific signal was detected using a CREB probe²³ in the pineal

gland at any time (Fig. 2b). Although CREM and CREB are closely related¹⁷, this dramatic difference in expression between the two genes confirms previous indications that CREM regulation is highly dynamic in endocrine systems^{19–21}. In addition, a more subtle day–night fluctuation of CREM expression was detected in the SCN, with elevated levels at daytime (not shown). To understand the significance of the CREM day–night rhythm within the physiology of the pineal gland, we investigated the mechanisms controlling its expression.

A time course of CREM expression in the pineal gland was defined by RNase protection analysis (Fig. 1a, b). There is a gradual increase in CREM expression from the light–dark transition, leading to a dramatic peak 7–9 h into the dark period (lanes 11–15). CREM transcripts decrease to a basal low level 3–4 h after the dark–light transition (lanes 16 and 5–10, Fig. 1b). The level of CREM expression in the pineal gland is greatly elevated at night time, being even higher than that in adult testis (compare lanes 2 and 14, 15 in Fig. 1b), a tissue with the highest abundance of CREM transcript¹⁹.

Accumulation of CREM messenger RNA in testis is controlled by the pituitary follicle stimulating hormone (FSH) and is due to a switch in the use of polyadenylation sites²⁰; this causes an increased transcript stability by virtue of the omission of mRNA destabilizers present in the 3' untranslated region of the gene^{18,20}. In contrast, accumulation of the pineal transcript is not accompanied by the switch in the use of polyadenylation sites (probe 2; Fig. 1a, c). Thus, control of mRNA stability by alternative polyadenylation does not apply in this case. Northern blot analysis confirms these findings (Fig. 1d, lanes 3, 4) and validates the dramatic day–night switch in CREM expression. Two major closely spaced bands of 2.1 kilobases (kb) and 1.7 kb hybridized with the CREM probe; we show that these two transcripts result from alternative use of the two DNA-binding domains¹⁸ by RNase protection analysis (Fig. 1b) and complementary DNA cloning (Fig. 3a, b). In addition, and consistent with the *in situ* results (Fig. 2b), CREB expression was very low and non-fluctuating, as monitored by northern analysis (Fig. 1d, lanes 5 and 6). To investigate the nature of the CREM pineal transcript, we used a probe (probe 3; Fig. 1a) specific for the N-terminal section of CREM (including the phosphorylation domain), common to all the previously characterized isoforms

‡ To whom correspondence should be addressed.

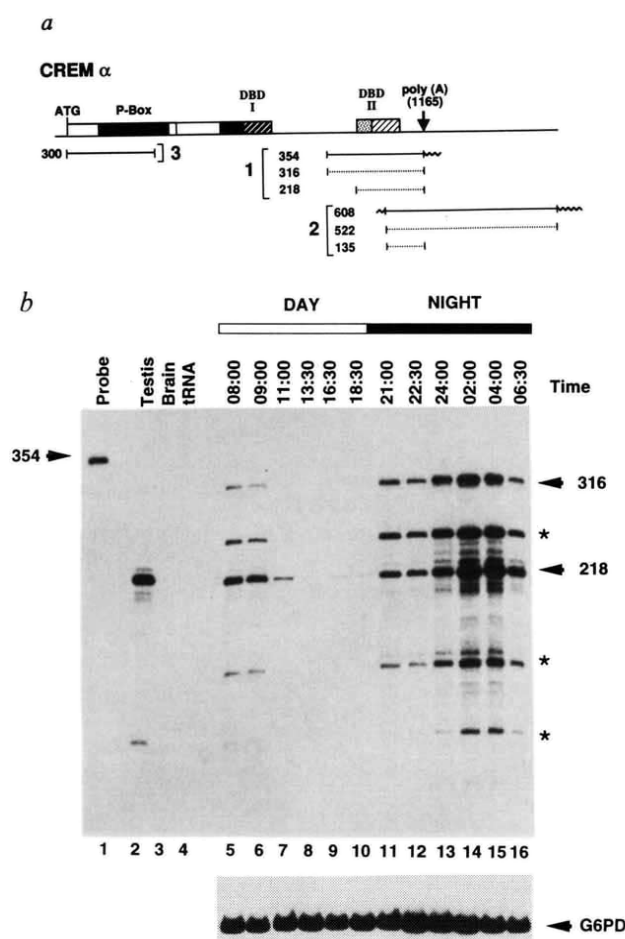
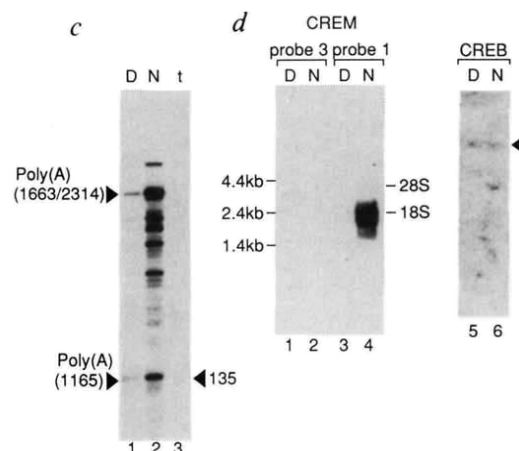


FIG. 1 Day-night analysis of CREM transcript in the pineal gland. *a*, Schematic representation of the CREM α cDNA¹⁸ showing the extent of probes used in CREM RNA analysis. The DNA-binding domains (DBDI and II) and the phosphorylation domain (P-box)¹⁸ are indicated. The solid bars below indicate the probes used. Dashed bars indicate the extent of partially protected probes obtained during the RNase protection analysis. Each probe is numbered and its size (in nucleotides (nt)) is indicated. *b*, RNase protection analysis using probe 1 of pineal gland total RNA at times indicated (2 μ g, lanes 5–16), or total RNA from mouse brain and testis (2 μ g, lanes 2–3). The sizes of bands, corresponding to the specific CREM-protected fragments are shown (nt); asterisks indicate bands generated by mismatches between rat RNA and the mouse

(α , β , γ and τ)^{18,19,24}. Importantly, this probe failed to hybridize with the pineal CREM transcripts (Fig. 1*d*, lanes 1, 2). Thus, the CREM night-time transcript is different from the other isoforms; indeed, as shown below, the cloning of the pineal CREM cDNA confirmed this prediction (Fig. 3).

An alternative intronic promoter

To characterize further the night-time-induced CREM transcript, we used a battery of exon-specific probes in a systematic northern *in situ* hybridization, RNase protection and reverse transcription-polymerase chain reaction (RT-PCR) analysis. Surprisingly, we established that all previously characterized exons 5' to the γ -exon were absent from the induced CREM transcript. This confirms the previous northern analysis (Fig. 1*d*), in which a probe incorporating the P-box and N-terminal exons failed to hybridize to the night-time pineal RNA. The 5' boundary of the γ -exon is defined by a consensus splicing acceptor site²⁵. Thus, we sought to identify the remaining exons constituting the 5' end. Adopting the rapid amplification of cDNA ends PCR (RACE-PCR) cloning strategy, we primed cDNA synthesis using an antisense oligonucleotide located



CREM probe²⁰. The bands of 316 nt and 218 nt correspond to CREM RNA incorporating the first or second DBD, respectively¹⁸. The ratio between these two bands does not fluctuate. Equal loading and quality of RNAs was confirmed by G6PD RT-PCR controls^{20,46}. Results have been reproduced in several independent experiments. *c*, Nocturnal CREM accumulation is not associated with a switch in polyadenylation sites (probe 2 in *a*). Bands arising from the use of specific polyadenylation sites are labelled with the equivalent positions of those sites in the mouse gene¹⁸. Additional bands are generated by mismatches between the mouse probe and rat RNA. The ratio between these bands does not fluctuate. *d*, Northern blot analysis of total pineal RNA (3 μ g) from animals at 12:00 (D) or 02:00 (N). Duplicate blots were hybridized with probe 3 (lanes 1, 2) or probe 1 (lanes 3, 4) and the full-length CREB probe (lanes 5, 6). The blots for CREM were exposed for 12 h, and the CREB blot was exposed for 3 days. Strikingly, no CREM hybridization is detected with probe 3, whereas a weak, non-fluctuating signal is visible for CREB.

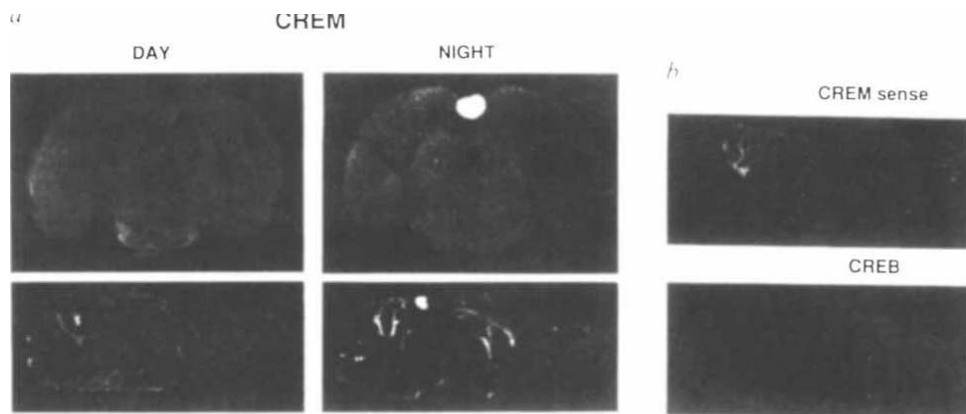
METHODS. RNA extraction, northern analysis and RNase protections were done as described^{18,20,45}. Riboprobe 1 is generated from p6N1 by *Nco*I digestion and *in vitro* transcription as described^{18,21} (see also Fig. 2). Riboprobe 2 is a 522 nt *Nsi*I-*Hind*III fragment straddling the polyadenylation site at position 1,165 of the mouse CREM cDNA sequence^{18,20}. Probe 3 is a 300 bp *Nco*I fragment containing the N-terminal and P-box regions¹⁸. Site positions correspond to those in the CREM α cDNA¹⁸. The CREB probe corresponds to the full-length rat cDNA²³. With each RNase protection probe we included a control of *Escherichia coli* transfer RNA and quantitation of RNA was determined by RT-PCR of G6PD transcript⁴⁵. The RT-PCR method and primer sequences have been previously reported²⁰. Each pineal RNA sample derives from 5–8 glands. The data presented are representative of a series of independent experiments.

downstream of the γ -exon (primer A in Fig. 3*b*). We generated a series of short, overlapping cDNA clones that define a novel cDNA (Fig. 3*a*). We found that an 82-base-pair (bp) sequence lies 5' of the γ -exon boundary and extends the CREM open reading frame upstream by eight amino acids to a consensus Kozak ATG codon²⁶. To confirm this result further, we screened a night-time pineal cDNA library and isolated 10 overlapping CREM clones. All these clones shared the same 5'-end sequence as the RACE-PCR clones. Considering the results from the northern analysis and cDNA cloning, we reasoned that the 5' end of the clones could represent a start of transcription. To test this hypothesis, we used a primer extension reaction using an antisense oligonucleotide primer specific for the novel exon (primer B in Fig. 3*b*). With night-time pineal RNA, we obtained a single major extension product corresponding to a start site of transcription 57 nucleotides upstream from the ATG codon (Fig. 3*c*, lanes 4–8). This site is consistent with the start points of most of the RACE clones and was confirmed by RNase protection analysis (Fig. 3*c*, lanes 1–3). Sequencing of the RACE-PCR and cDNA clones revealed differential splicing of the two alternative DNA binding domains and of the γ -exon¹⁸, as previ-

FIG. 2 CREM expression is elevated at night in the pineal gland. *a*, *In situ* hybridization of mid-sagittal and coronal brain sections with an anti-sense CREM-specific riboprobe. Rats maintained in 12 h light/12 h dark conditions (light on 07:00) were killed at consecutive timepoints. Representative sections from 12:00 (day) and 02:00 (night) are shown. An intense hybridization signal is present in the pineal gland at night. *b*, Sections from rats killed at 02:00 were also hybridized with a sense CREM riboprobe or an antisense riboprobe generated from a full-length CREB cDNA fragment.

METHODS. *In situ* hybridization techniques are described elsewhere⁴⁵.

³⁵S-labelled CREM sense and antisense riboprobes were generated by *in vitro* transcription from p6N1, a 601-bp *Hpa*II CREM α subclone spanning the second DNA-binding domain^{18,21} (probe 1, see Fig. 1a) and linearized by *Xho*I or *Nco*I, respectively. Sections were exposed to X-ray film for 7–10 days. As control, adjacent sections were pretreated with



RNase A before hybridization (not shown). All animals used in this work were male Wistar rats obtained from Iffa-Credo (France). These results have been reproducibly obtained from tissue samples of several rats in independent experiments.

ously predicted by the northern analysis (Fig. 1d, lane 4); this phenomenon was previously described for the CREM α and β isoforms²⁴. Thus, even for this novel CREM isoform the hallmark of the alternative use of DNA binding domains is conserved^{18,24}.

We next investigated whether the start point of transcription of the pineal transcript is the same as for the previously described isoforms (upstream promoter, P1 in Fig. 3b). By screening an overlapping series of phage clones encompassing the entire CREM gene, we located the 5' terminal exon within the 10-kb intron lies between the 3' glutamine-rich domain exon (Q2; ref. 19) and the γ -exon. Thus, the pineal transcript is generated by the use of an alternative internal promoter, which we termed P2. The relative position of P2 with respect to P1 is depicted in Fig. 3b. A comprehensive study of the structure and function of this promoter is presented elsewhere²⁵.

ICER is a remarkably small factor

The transcript generated from P2 is constituted by the C-terminal segment of CREM. The predicted open reading frame encodes a small protein of 120 amino acids with an expected M_r of 13.4K (Fig. 3a). The presence of this protein *in vivo* is confirmed by our analysis with CREM-specific antibodies²⁷ (see later and Fig. 4a). This protein, compared with the previously described CREM isoforms^{18,19}, essentially consists of only the DNA-binding domain, which is constituted by the leucine-zipper and basic region¹⁸. We have termed this protein ICER (inducible cAMP early regulator). The structure of ICER is suggestive of its function and makes it one of the smallest transcription factors ever described.

We cloned the ICER coding sequence into both a prokaryotic and an SV40-based eukaryotic expression vector¹⁸. ICER protein generated from transiently transfected cultured cells and from bacteria with CREM-specific antibodies was analysed by western blot. The comigration on SDS-PAGE (Fig. 4a; compare lanes 2 and 4) confirmed the use of the canonical Kozak ATG²⁵. This band is absent in extracts from cells transfected with a control plasmid (lane 3). Expression of ICER-CREM in night- and day-time pineal gland extracts was also analysed. A clear difference in ICER levels is evident, agreeing with the RNA analysis (lanes 5 and 6). Importantly, the CREM-immunoreactive band in pineal extracts precisely comigrates with ICER expressed in both eukaryotic and prokaryotic systems²⁸ (compare lanes 2, 4 and 6). A fainter, slightly higher- M_r band is also visible in the extract from night-time pineal gland. This product represents the ICER form that includes the alternatively spliced γ -domain as shown by the cDNA cloning (see previous cDNA

cloning section and also ref. 25). Other higher- M_r immunoreactive bands are detected and correspond to P1-generated CREM isoforms which do not fluctuate between day and night (only CREM γ is visible in lanes 5–7).

ICER represses cAMP-induced transcription

The presence of an intact DNA-binding domain predicted that ICER would be able to bind specifically to a consensus CRE element¹⁸. We show that ICER generated in both bacteria and transfected cells binds efficiently to CREs (Fig. 4b, lanes 1–9). Consistent with previous analysis of the other CREM isoforms, ICER fails to bind to AP-1 or Sp1 sites¹⁸ (lanes 10–13). Importantly, ICER is able to heterodimerize with CREM τ (Fig. 4b, lanes 14–16), as well as with the other CREM proteins and with CREB (not shown). We next addressed the function of ICER as a transcriptional regulator. To use the most appropriate system, we dissociated pineal glands into culture dishes to transfect primary pinealocytes¹¹. An attractive feature of this system is that pinealocytes constitute 90% of the cell population. In addition, we transfected JEG-3 cells. Cells were transfected using an extensive range of reporter plasmids carrying individual CRE elements or cAMP-inducible promoter fragments. In all cases, treatment with forskolin or cotransfection of the catalytic subunit of PKA²⁹ causes a significant induction of expression from these reporters (Fig. 4c). Here we show that coexpression with ICER completely abolishes cAMP-mediated induction in both primary pinealocytes and JEG-3 cells. Repression is specific as tested in a GAL4/VP16 transactivation system (not shown, and ref. 30). Interestingly, ICER-mediated repression is obtained at substoichiometric concentrations (not shown), similar to the previously described CREM antagonists^{18,22}. It should be noted that ICER is much more potent than CREM β (Fig. 4c).

Clock-distal elements control ICER expression

Given the striking day-night fluctuation of ICER expression and its role as a transcriptional repressor, we sought to dissect the physiological mechanisms controlling its rhythmic expression in the pineal gland. Animals maintained in constant darkness (DD) were killed at different times. The temporal switch in ICER expression is conserved under DD conditions (Fig. 5a; compare lanes 1 and 2 with 5 and 6, where LD indicates a normal lighting regimen), which demonstrates that the ICER rhythm is driven by an endogenous clock mechanism. In contrast, under constant light treatment (LL) no ICER switch is observed (Fig. 5a; compare lanes 3, 4 with 5, 6). LL conditions are known to dissociate the coordination of the clock output³¹ and to suppress melatonin synthesis^{32,33}. In addition, a light pulse at night, known to

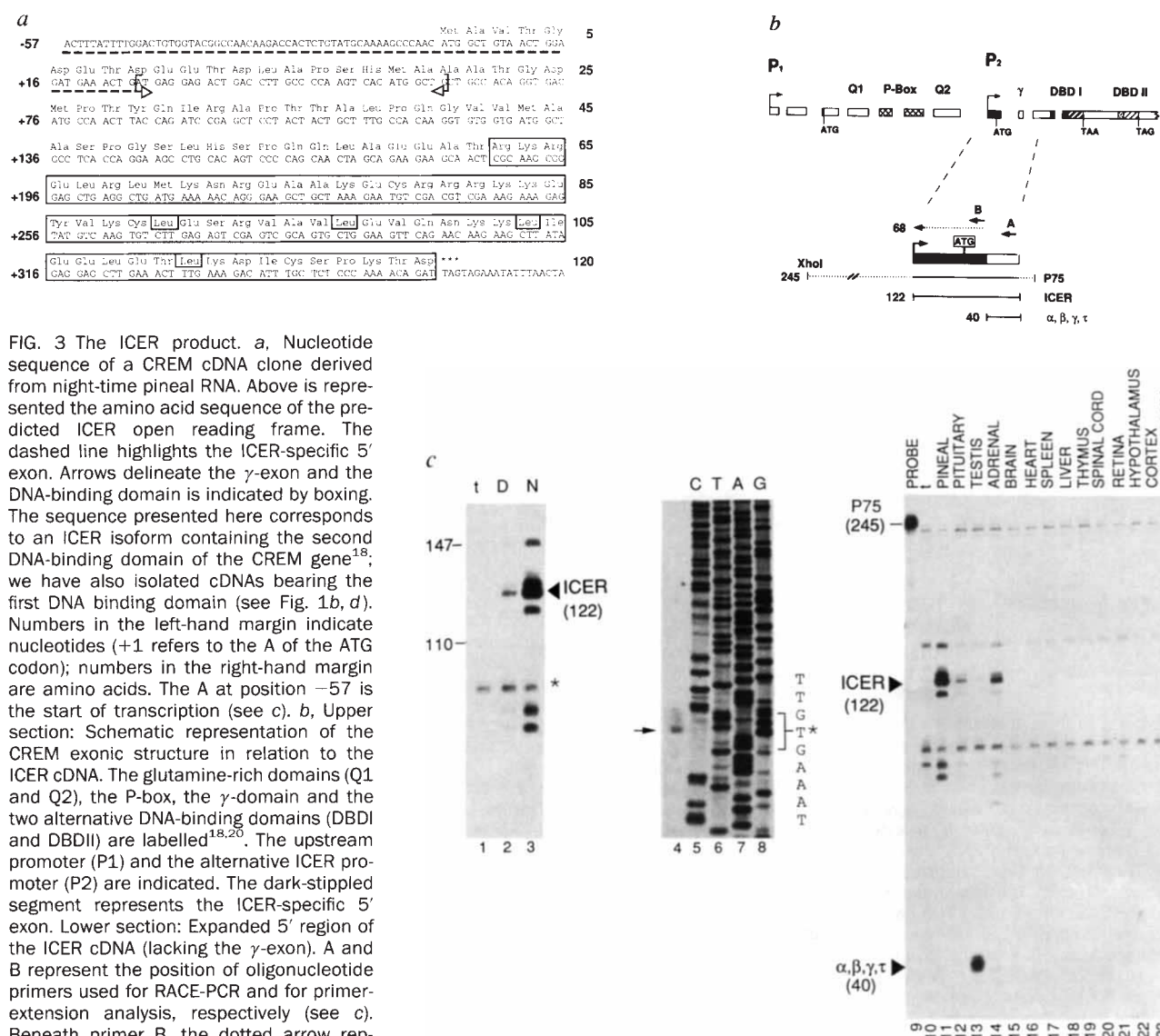


FIG. 3 The ICER product. **a**, Nucleotide sequence of a CREM cDNA clone derived from night-time pineal RNA. Above is represented the amino acid sequence of the predicted ICER open reading frame. The dashed line highlights the ICER-specific 5' exon. Arrows delineate the γ -exon and the DNA-binding domain is indicated by boxing. The sequence presented here corresponds to an ICER isoform containing the second DNA-binding domain of the CREM gene¹⁸; we have also isolated cDNAs bearing the first DNA binding domain (see Fig. 1b, d). Numbers in the left-hand margin indicate nucleotides (+1 refers to the A of the ATG codon); numbers in the right-hand margin are amino acids. The A at position -57 is the start of transcription (see c). **b**, Upper section: Schematic representation of the CREM exonic structure in relation to the ICER cDNA. The glutamine-rich domains (Q1 and Q2), the P-box, the γ -domain and the two alternative DNA-binding domains (DBD I and DBD II) are labelled^{18,20}. The upstream promoter (P1) and the alternative ICER promoter (P2) are indicated. The dark-stippled segment represents the ICER-specific 5' exon. Lower section: Expanded 5' region of the ICER cDNA (lacking the γ -exon). A and B represent the position of oligonucleotide primers used for RACE-PCR and for primer-extension analysis, respectively (see c). Beneath primer B, the dotted arrow represents the primer-extended product. The extent of the probe P75 used for RNase protection assays is shown together with the protected fragments for ICER and the previously described CREM isoforms α , β , γ and τ ^{18,20}. For each fragment, the size is indicated (nt). **c**, Lanes 1–3; RNase protection analysis of day (lane 2) and night (lane 3) pineal RNA (2 μ g) using the P75 probe. The size of the protected fragment indicated by an arrowhead (ICER, 122 nt) confirms the major start of transcription for the promoter P2 and the day–night fluctuation of CREM expression. The asterisk indicates a non-specific band; t, tRNA control; the position of size markers is indicated in the left-hand margin. Lanes 4–8; primer extension analysis using primer B (see b) with night-time pineal RNA (10 μ g total RNA). The major start site is indicated by an arrowhead. The sequencing reactions of an overlapping genomic fragment using primer B were used to pinpoint the start of transcription (lanes 5–8 and right-hand margin for corre-

sponding sequence). Lanes 9–25; RNase protection analysis with probe P75 of ICER expression in a panel of rat tissues; lane 9, probe alone (P75); lane 10, tRNA control. The tissue for each RNA preparation is indicated (lanes 11–25). Arrows in the left margin indicate bands protected by transcripts generated from the P1 (α , β , γ , τ) or P2 (ICER) promoters. Size markers (bp) and nonspecific bands (asterisk) are indicated in the right margin.

METHODS. The night-time pineal cDNA library used to clone ICER-CREM was constructed in λ -Zap (Stratagene). RNA techniques were as described in Fig. 1 legend and in refs 18 and 20. Primer extension was as described⁴⁶. Sequence of primers A and B were 5'-GGGGACTGTGACAGGCTCTCT-3' and 5'-GTTACAGCCATGTTGGGCTTTTGC-3', respectively. RACE-PCR was done using a standard procedure⁴⁸.

decrease melatonin production³⁴, also greatly blocked ICER induction (lanes 7–9).

We next tried to define the signal whereby ICER circadian expression is driven by clock-distal elements. At night, post-ganglionic fibres originating from the superior cervical ganglia (SCG) release noradrenaline^{35,36}, which in turn regulates melatonin synthesis through β -adrenergic receptors⁹. Injection of the β -adrenergic antagonist propranolol before the onset of darkness blocks accumulation of ICER mRNA (Fig. 5b, lanes 1 and 2) and melatonin synthesis³⁷. The same effect was obtained by chronic denervation of pineal glands with bilateral removal of the SCG (SCGX, lanes 3–6). To establish unambiguously the

role of noradrenaline, we show that the β -adrenergic agonist isoproterenol (Iso) induces ICER expression in ganglionectomized animals (lanes 9–12), concomitantly with melatonin synthesis³⁷.

Isoproterenol injections in intact rats stimulate melatonin synthesis independently of the time of the day³⁷. Unexpectedly, no ICER induction was observed on isoproterenol administration during daytime (Fig. 6a, lanes 1–3), even 6 h after injection (not shown). This is in contrast with SCGX animals (Fig. 5b, lanes 9–12), where isoproterenol induces ICER expression during both day and night. To investigate further this surprising disparity and to circumvent the well documented *in vivo* fluctuations of

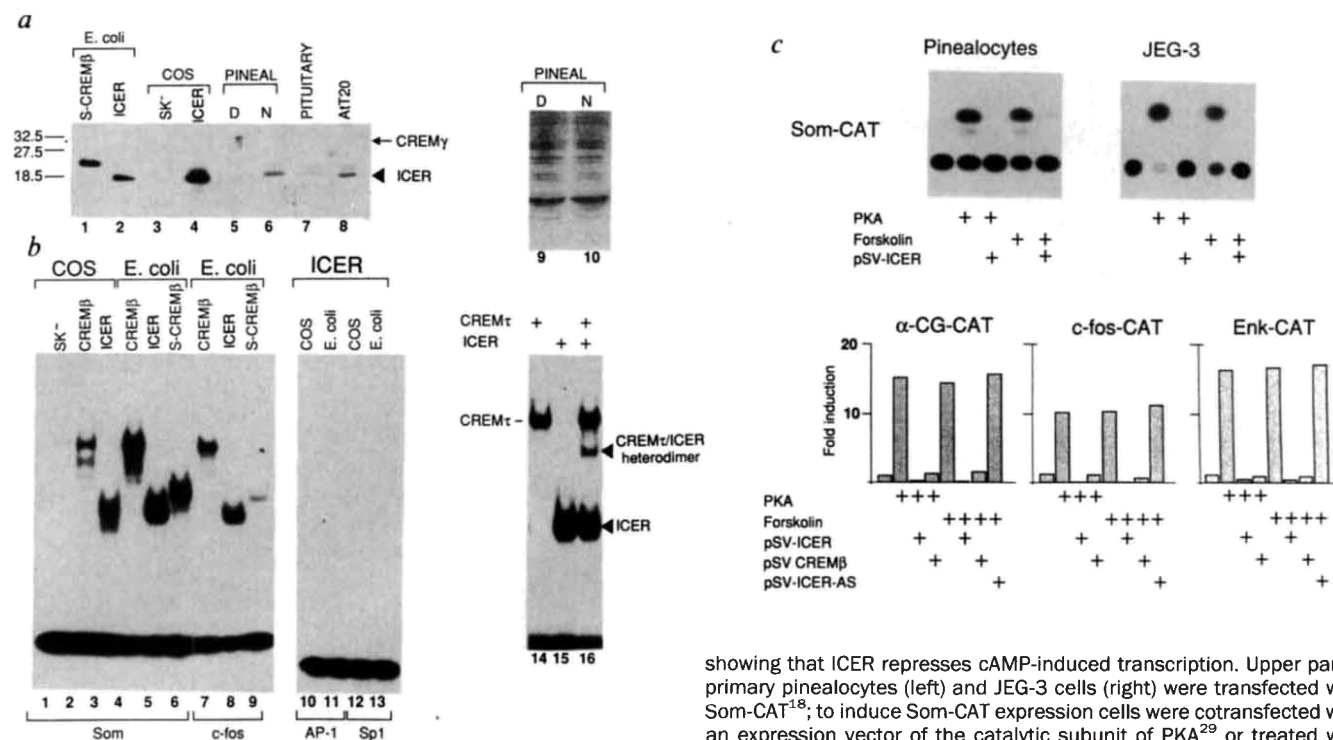


FIG. 4 ICER is a repressor of cAMP-induced transcription. **a**, Western blot analysis of ICER protein. Lane 1, S-CREB β ²⁴ and lane 2, ICER-CREB generated in bacteria (*E. coli*); lanes 3 and 4, extracts from COS cells transfected with control plasmid SK⁻ or pSV-ICER (ICER) respectively (pSV-ICER contains the ICER coding sequence in the pSG5 expression vector¹⁸). Lanes 5–6, pineal gland extracts from rats killed during the early light-phase (lane 5: D) and the middle of the dark phase (lane 6: N); lane 7, extract from rat pituitary gland and lane 8, extract from AtT-20 corticotroph cells. The arrowhead indicates the position of ICER, which comigrates in all samples. The positions of size markers (*M*, in thousands) are labelled in the left-hand margin. ICER-CREB, like the other CREB products, runs with a slower mobility with respect to the predicted size on SDS-PAGE²⁴. Equal loading of the pineal extracts was determined by Coomassie blue staining of the protein gel (lanes 9–10). **b**, DNA binding gel-shift analysis of ICER. Proteins generated in either COS-transfected cells or in bacteria (*E. coli*) were tested for binding to various CREs. Here only a representative analysis for the somatostatin¹⁸ (lanes 1–6) and *c-fos*³⁰ (lanes 7–9) elements is shown. Migration of complexes containing ICER is faster than with S-CREB β ²⁴, as predicted from the size difference (compare lanes 5 and 6 and lanes 8 and 9). A control COS extract was obtained from cells transfected with SK⁻ (lane 1). Competition experiments confirmed the specificity of binding (not shown). No binding of bacterial (*E. coli*) or COS cell generated ICER was detected in the same conditions to AP-1 and Sp1 sites (lanes 10–13). ICER efficiently heterodimerizes with CREB τ ; the complex including the heterodimer is indicated (lanes 14–16). **c**, Transfection experiments

showing that ICER represses cAMP-induced transcription. Upper panel: primary pinealocytes (left) and JEG-3 cells (right) were transfected with Som-CAT¹⁸; to induce Som-CAT expression cells were cotransfected with an expression vector of the catalytic subunit of PKA²⁹ or treated with forskolin for 3 h³⁰, as indicated. Cotransfection with the ICER expression vector (pSV-ICER) in both cell types represses activation. Lower panel: summary of transfection experiments with other reporter vectors containing either isolated CRE elements or complete cAMP-inducible promoters^{18,29,30}. These reporters were cotransfected with expression vectors for the catalytic subunit of the PKA²⁹, CREB β ¹⁸ or ICER. ICER represses expression whether it is induced by cotransfection with PKA or treatment with forskolin. No effect is obtained with an antisense ICER expression vector (pSV-ICER-AS). Repression by ICER appears to be at least 5–10-fold stronger than by CREB β . *c-fos*-CAT contains a segment of the *c-fos* human promoter³⁰; Enk-CAT contains a segment of the enkephalin promoter³⁰; α -CG-CAT contains the CREs from the α -chorionic gonadotropin gene³⁰. **METHODS.** The antibodies used in the western analysis²⁷ were raised against the CREB τ bacterially expressed protein. These antibodies are specific for CREB, and do not crossreact with CREB (V. Delmas *et al.*, manuscript in preparation, and ref. 27). DNA binding and dimerization assays were done as described^{18,24}. Primary pinealocyte cultures were prepared as described elsewhere¹¹. Transfection assays were done by the calcium phosphate co-precipitation technique. Typically per 10-cm dish of JEG-3 cells, we transfected 2 μ g of the reporter, 400 ng PKA and 300 ng of the various CREB expression vectors. For the primary pinealocytes, 4 μ g reporter, 1 μ g PKA and 4 μ g CREB expression vectors were used per dish. Forskolin treatment, CAT assays and quantification methods are described elsewhere¹⁸. The results shown here are representative of a series of several independent experiments. Variation in the effects was less than 10%.

pineal β -adrenergic receptor sensitivity³⁸, we used the non-hydrolysable cAMP analogue, dBcAMP. We stimulated explanted pineal glands in an *in vitro* superfusion system³⁹. Glands taken from animals immediately before the onset of darkness readily showed induction of ICER (Fig. 6c, lanes 1–8) and melatonin synthesis. In contrast, the same treatment failed to induce ICER expression in glands taken from animals immediately after the onset of light (Fig. 6b, lanes 1–8), although melatonin synthesis was stimulated as expected. Thus, ICER represents the first case of day-night disparity in cAMP-inducibility. In conclusion, signals from the SCN dictate the competence of ICER inducibility.

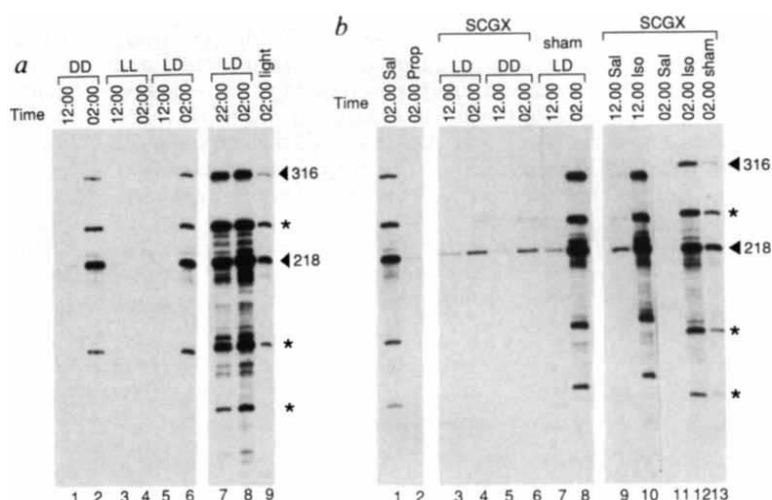
ICER is the major CREM transcript

Given the high level and its profound rhythmic regulation in rat pineal gland, we reasoned that ICER may be important in other

neural and endocrine tissues. We thus analysed ICER expression in a range of rat tissues by RNase protection assay. A probe was generated from the 5' end of the ICER transcript (P75 in Fig. 3b), which scores for both ICER and the transcripts initiated at the upstream P1 promoter. Although the overall amount of CREM message varies considerably between different tissues, in each sample the major protected fragment corresponds to ICER, with the exception of testis (Fig. 3c). Indeed, as predicted from previous studies¹⁹, the major protected fragment in testis corresponds to transcripts initiated at the P1 promoter.

A striking observation from this RNase protection analysis is a tissue-specific pattern of ICER expression. ICER is expressed at high levels predominantly in tissues of neuroendocrine origin (lanes 11, 12, 14), namely the pineal, pituitary and adrenal glands (also confirmed by *in situ* hybridization; results not shown).

FIG. 5 CREM expression is under the control of the endogenous clock. *a*, RNase protection of pineal RNA using probe 1 (Fig. 1*a*) from animals maintained in different lighting conditions (DD: constant darkness; LL: constant light; LD: 12 h light:12 h dark) (lanes 1–6). Animals were killed during daytime (12:00) or night time (02:00). In addition, animals kept under LD (lanes 7, 8) show a dramatic reduction in CREM expression when exposed to 4 h bright white light (lane 9), starting at 22:00. *b*, Injection of rats with the β -adrenergic antagonist propranolol (Prop; 20 mg kg⁻¹ body weight) before the onset of darkness (18:45) and killed at 02:00 caused a dramatic reduction in CREM expression (lane 2); injection of a saline (Sal) solution is shown as control (lane 1). Lanes 3–6, CREM expression at day and night in animals in which, 10 days before experimentation, the superior cervical ganglion was removed bilaterally³⁷ (SCGX). One group of SCGX animals was transferred one day before experimentation into constant darkness (DD, lanes 5, 6). Rats were killed at the times indicated. One group of animals was sham-operated (lanes 7, 8). Lanes 9–12, effect of injection of the β -adrenergic agonist isoproterenol (Iso; 5 mg kg⁻¹ body weight) on CREM expression. SCGX animals received the injection 2 h before indicated killing time. Lane 13, pineal RNA from a control group of intact rats killed at 02:00. The higher level of CREM expression in the injected SCGX animals (compare lanes 12 and 13) is likely to be due to supersensitivity of the pineal β -adrenergic receptors^{37,38}. The CREM isoform detected in these experiments corre-



sponds to ICER as confirmed using probe P75 (Fig. 3*b*).

METHODS. Experiments are described in the legend to Fig. 1, and were done in duplicate; 4–6 animals were used for each point. Melatonin concentrations were assayed from trunk blood³⁹.

ICER expression in the pituitary does not fluctuate (not shown). In all other neuronal and non-endocrine tissues tested (lanes 15–25), ICER is expressed in uniformly low amounts. The presence of the ICER protein was also confirmed in other cell types (Fig. 4*a*, lanes 7, 8). Note that the ICER transcript was not detected in previous analyses as a result of the strategies used. Indeed, primers and probes used previously in RT-PCR and RNase protection^{18,19,24} studies could not have discriminated products from the alternative P2 promoter.

Discussion

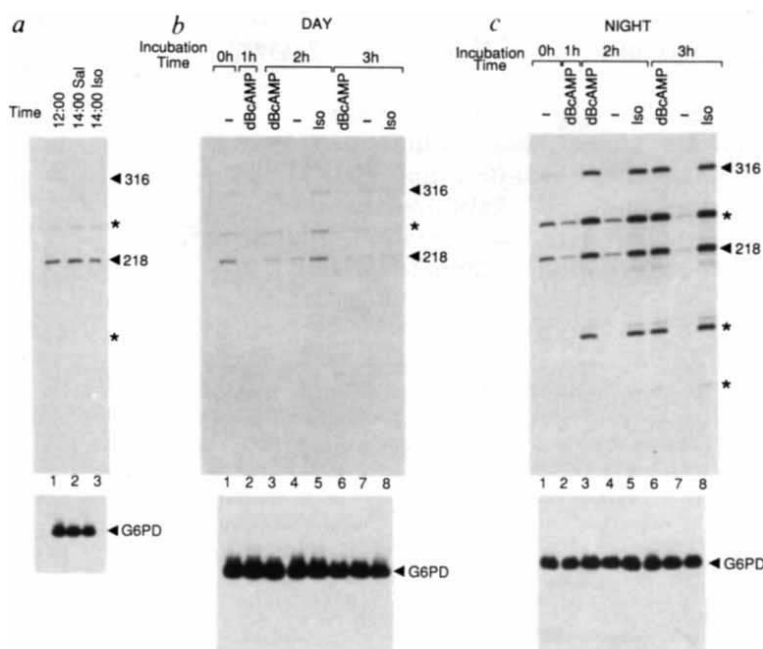
We have previously shown that CREM expression is hormonally regulated during spermatogenesis²⁰. Here we describe a new facet of CREM regulation within the neuroendocrine reproductive axis. In the pineal gland a striking circadian fluctuation in CREM mRNA is under neural control. We have shown that this phenomenon depends on adrenergic stimulation of the cAMP

pathway. Thus, environmental light–dark transitions, together with the endogenous clock, modulate the expression of this transcriptional regulator. The pineal CREM isoform, ICER, encodes a powerful repressor of cAMP-induced transcription. The structural organization of the CREM gene, which allows the production of ICER, is unique among the genes of the CRE-binding protein class^{15–17}. By the use of an alternative, intronic promoter, a very small transcriptional regulator is generated which has only the DNA-binding function^{15,17}. The omission of the phosphorylation domain in ICER indicates that post-translational modifications do not guide its function. Thus, the intracellular level of ICER should be the sole determinant of its activity.

Circadian regulation of gene expression has been previously documented^{4,40}, although the molecular and physiological mechanisms underlying these phenomena remain unclear. In this regard, many examples of cyclic expression are related to genes

FIG. 6 CREM inducibility in pineal glands depends on the time of day. *a*, Lanes 1–3, CREM is not induced after injection of Iso (5 mg kg⁻¹ body weight) 2 h before killing time (14:00). *b*, *c*, Important differences in CREM inducibility on application of dibutyryladenosine 3',5'-cyclic monophosphate (dBcAMP, 10⁻³ M) or Iso (10⁻⁶ M) to *in vitro* pineal gland cultures. Glands were explanted in early morning (*B*, day, lanes 1–8; 07:30) or late day (*C*, night, lanes 1–8; 18:30) and subsequently maintained in a dynamic superfusion system³⁹. In all experiments the use of probe P75 confirmed that the transcript detected corresponded to ICER.

METHODS. For the *in vitro* experiments, explanted glands were preincubated for 5 h before stimulation to eliminate initial norepinephrine release from sympathetic nerve endings. RNase protection analysis has been described^{18,20}. Equal loading and quality of RNAs was confirmed by G6PD RT-PCR controls⁴⁵. Results have been reproduced in independent experiments. Melatonin concentration for *in vitro* experiments was determined by radioimmunoassay from perfusate collected in 20-min bins as described³⁹.



of the early response class⁴⁰. Note that among the previously described regulators that show circadian expression (DBP⁴¹ and *per*^{42,43}), ICER is the first to be demonstrated to be a repressor. CREB phosphorylation is also modulated in the SCN during the circadian cycle⁴⁴. In this respect, the large fluctuation in the level of the ICER repressor might be the complementary facet of a fulcral transduction mechanism conveying circadian rhythmicity.

It is tempting to speculate on the significance of circadian regulation of a repressor in the control of pineal gland physiology. These observations constitute the first example, to our

knowledge, where concomitance of melatonin synthesis and inducibility of a specific gene can be dissociated. A fascinating model to explain why CREM is refractory to induction during daytime (Fig. 6b) implicates a cyclic block of the ICER regulatory region by a negative factor. Importantly we have shown that ICER negatively autoregulates its own promoter in cultured cells²⁵, suggesting that the same mechanism may operate in the pineal gland. The findings reported here directly implicate CREM as a major player in governing neuroendocrine mechanisms regulated by the cAMP-dependent signal transduction pathway. □

Received 11 May; accepted 16 August 1993.

1. Aschoff, J. (ed.) in *Handbook of Behavioural Neurobiology* Vol. 4, 1–10 (Plenum, New York, 1981).
2. Moore-Ede, M. C., Sulzmann, F. M. & Fuller, C. A. *The Clocks that Time Us* (Harvard Univ. Press, London, 1982).
3. Darwin, C. *The Power of Movement in Plants* (Murray, London, 1880).
4. Rosbash, M. & Hall, J. C. *Neuron* **3**, 387–398 (1989).
5. Tamarkin, L., Baird, C. J. & Almeida, O. F. X. *Science* **227**, 714–720 (1985).
6. Reiter, R. J. *Trends Endocr. Met.* **1**, 13–19 (1991).
7. Klein, D. C., Moore, R. Y. & Reppert, S. M. (eds) *Suprachiasmatic Nucleus: The Mind's Clock* (Oxford Univ. Press, Oxford, 1991).
8. Moore, R. Y. *Federation Proc.* **42**, 2783–2789 (1983).
9. Klein, D. C. in *Photoperiodism, Melatonin and the Pineal Gland* (Pitman, London 1985).
10. Sudgen, D., Vanecek, J., Klein, D. C. & Thomas, T. P. *Nature* **314**, 359–361 (1985).
11. Vanecek, J., Sudgen, D., Weller, J. & Klein, D. C. *Endocrinology* **116**, 2167–2173 (1985).
12. Hoffmann, K. in *Handbook of Behavioural Neurobiology* Vol. 4 (ed. Aschoff, J.) 449–473 (Plenum, New York, 1981).
13. Pévet, P. *Reprod. Nutr. Dev.* **28**, 575–578 (1988).
14. Carter, D. S. & Goldman, B. D. *Endocrinology* **113**, 1261–1267 (1983).
15. Habener, J. *Molec. Endocr.* **4**, 1087–1094 (1990).
16. Ziff, E. B. *Trends Genet.* **6**, 69–72 (1990).
17. deGroot, R. P. & Sassone-Corsi, P. *Molec. Endocr.* **7**, 145–153 (1993).
18. Foulkes, N. S., Borrelli, E. & Sassone-Corsi, P. *Cell* **64**, 739–749 (1991).
19. Foulkes, N. S., Mellström, B., Benusiglio, E. & Sassone-Corsi, P. *Nature* **355**, 80–84 (1992).
20. Foulkes, N. S., Schlotter, F., Pévet, P. & Sassone-Corsi, P. *Nature* **362**, 264–267 (1993).
21. Mellström, B., Naranjo, J. R., Foulkes, N. S., Lafarga, M. & Sassone-Corsi, P. *Neuron* **10**, 655–665 (1993).
22. Foulkes, N. S. & Sassone-Corsi, P. *Cell* **68**, 411–414 (1992).
23. Gonzales, G. A. *et al. Nature* **337**, 749–752 (1989).
24. Laoidé, B. M., Foulkes, N. S., Schlotter, F. & Sassone-Corsi, P. *EMBO J.* **12**, 1179–1191 (1993).
25. Molina, C. A., Foulkes, N. S., Lali, E. & Sassone-Corsi, P. *Cell* (in the press).
26. Kozak, M. J. *Cell Biol.* **108**, 229–241 (1989).
27. Delmas, V. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 4226–4230 (1992).
28. Studier, F. W., Rosenberg, A. M., Dunn, J. J. & Dubendorff, J. W. *Meth. Enzym.* **185**, 60–89 (1990).

29. Mellon, P. M., Clegg, C. H., Correll, L. A. & McKnight, G. S. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4887–4891 (1989).
30. Foulkes, N. S., Laoidé, B. M., Schlotter, F. & Sassone-Corsi, P. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5448–5452 (1991).
31. Meijer, G. H. in *Suprachiasmatic Nucleus: The Mind's Clock* (eds Klein, D. C., Moore, R. Y. & Reppert, S. M.) (Oxford Univ. Press, Oxford, 1991).
32. Tamarkin, L., Reppert, S. M. & Klein, D. C. *Endocrinology* **104**, 385–389 (1979).
33. Perlow, M. J., Reppert, S. M., Tamarkin, L., Wyatt, R. J. & Klein, D. C. *Brain Res.* **182**, 211–216 (1980).
34. Ilnerova, H., Vanecek, J., Krecek, J., Wetterberg, L. & Sääf, J. J. *Neurochem.* **32**, 673–681 (1979).
35. Brownstein, M. & Axelrod, J. *Science* **184**, 163–165 (1974).
36. Craft, C. M., Morgan, W. W. & Reiter, R. J. *Neuroendocrinology* **38**, 193–198 (1984).
37. Axelrod, J. *Science* **184**, 1341–1348 (1974).
38. Romero, J. & Axelrod, J. *Science* **184**, 1091–1092 (1974).
39. Simonneaux, V. *et al. J. Pineal Res.* **7**, 63–83 (1989).
40. Takahashi, J. S. *Curr. Opin. Genet. Dev.* **3**, 301–309 (1993).
41. Wuarin, J. & Schibler, U. *Cell* **63**, 1257–1266 (1990).
42. Takahashi, J. S. *Science* **258**, 238–240 (1992).
43. Hardin, P. E., Hall, J. C. & Rosbash, M. *Nature* **343**, 536–540 (1990).
44. Ginty, D. D. *et al. Science* **260**, 238–241 (1993).
45. Stehle, J. H. *et al. Molec. Endocr.* **6**, 384–393 (1992).
46. Vulliamy, T. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 5171–5175 (1988).
47. Williams, J. G. & Mason, P. J. in *Nucleic Acids Hybridization* (eds Hames, B. D. & Higgins, S. J.) 139–160 (IRL, Oxford, 1987).
48. Frohman, M. A. in *PCR Protocols* (eds Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J.) 28–38 (Academic, San Diego, 1990).

ACKNOWLEDGEMENTS. J.H.S. and N.S.F. should be considered as equal first authors. We thank E. Borrelli, M. Masson-Pévet, A. Giangrande, D. C. Klein, E. Lalli, J. P. Montmayeur and S. Schotten-Stehle for discussions, C. Craft for the pineal cDNA library, F. Schlotter, D. Bonn, L. Richie, M. X. Little, M. Rikkers, D. Chang Lam, H. Dazs, B. Boulay, M. Christian and C. Werlé for help. J. H. S. is supported by a European Economic Community fellowship; N.S.F. by the Association pour la Recherche contre le Cancer; C.A.M. by a National Science Foundation fellowship. This work was supported by CNRS, INSERM, FRM, CHUR, ARC and Rhône-Poulenc Rorer.

LETTERS TO NATURE

An image of the Sunyaev-Zel'dovich effect

Michael Jones*, Richard Saunders*, Paul Alexander*, Mark Birkinshaw†, Nick Dillon*, Keith Grainge*, Stephen Hancock*, Anthony Lasenby*, Dominic Lefebvre*, Guy Pooley*, Paul Scott*, David Titterton* & Donald Wilson*

* Mullard Radio Astronomy Observatory, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK

THE Sunyaev-Zel'dovich effect¹ is a distortion imposed on the 2.7-K cosmic microwave background radiation when the microwave photons are scattered by the hot gas in galaxy clusters. At radio wavelengths it appears as a decrement (of about 0.5 mK) in the brightness temperature of the background radiation. Measurements of this effect can provide information about the physics of galaxy clusters, and can also potentially be used to determine the

Hubble constant, by constraining the size of the cluster along the line of sight. Successful detections have been made with single-dish radio telescopes^{2,3}, from which one-dimensional profiles of the temperature decrement are constructed. But these observations are susceptible to several systematic errors, such as corruption of the measured signal by radio sources located close to the line of sight. To circumvent these problems, we have obtained a two-dimensional image, using an interferometric (rather than single-dish) approach, of structure in the microwave background in the vicinity of the galaxy cluster Abell 2218. Our measurements, combined with X-ray observations of the scattering gas in the same cluster, support a low value of the Hubble constant.

In the Rayleigh-Jeans region of the cosmic microwave background radiation (CMBR) spectrum (that is, at frequencies <200 GHz), the decrease ΔT in the brightness temperature of the CMBR owing to the Sunyaev-Zel'dovich (S-Z) effect is given by

$$\Delta T = -2T_{\text{CMBR}} \int \sigma_T n_e \frac{k_B T_e}{m_e c^2} dl$$

where n_e and T_e are the electron number density and temperature in the cluster gas, σ_T , m_e , k_B and c are respectively the Thomson cross-section, the electron mass, Boltzmann's constant and the speed of light, and the integral is along the line of sight. Thus ΔT depends only on the line integral of the gas pressure through the cluster. This has two immediate consequences. First, as the

† Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts, USA.

‡ Present Address: Herchel Smith laboratory for Medicinal Chemistry, University Forvie Site, Robinson Way, Cambridge, UK.

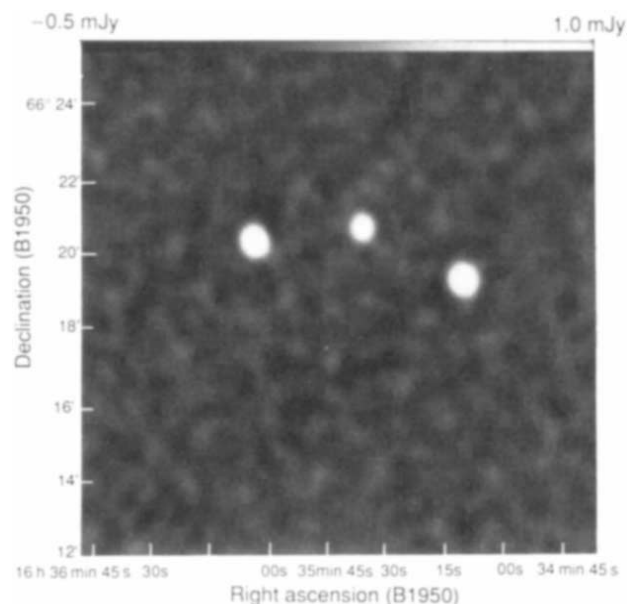


FIG. 1 A map derived from 27 twelve-hour observations of Abell 2218, seven in an array providing baselines of 18, 18, 36, 36, 36, 54, 72, 72, 90 and 108 m, the remainder in an array providing baselines of 18, 27, 27, 36, 45, 45, 63, 72, 81 and 108 m. Only baselines of 36 m and longer were used in this map. The unresolved source 1642+690 was observed as a phase calibrator every 20 min during the observations, and the flux scale established by observations of 3C286 (assumed to have a total flux density of 3.50 Jy in Stokes parameter $I+Q$) at the end of each observation. The pointing centre used was the centroid of the Rosat X-ray image, 16 h 35 min 42.0 s, +66° 18' 46" (G. Stewart, private communication). The phase or fringe-stopping centre was placed 3 arcmin south of this, as certain types of instrumental problem might generate spurious signals at the phase centre. (In fact there is no evidence of any artefacts at the phase centre in the final maps.) The map has been CLEANed and restored with a beam of 30×34 arcsec position angle (PA) = 2°. The grey-scale at the top of the figure runs from -0.5 to +1.0 mJy. All coordinates are epoch B1950.

size of the effect is independent of the distance to the cluster, it can in principle be detected out to very large redshifts, well beyond the range of optical telescopes (which detect the cluster galaxies) or X-ray telescopes (which can detect thermal emission from the cluster gas). Second, for relatively nearby clusters (redshift $z \leq 0.3$) which can be detected in X-rays, one can solve for the physical size of the cluster along the line of sight: this is done by combining X-ray *bremsstrahlung* measurements (which are proportional to $\int n_e^2 T_e^{1/2} dl$) and X-ray spectral measurements (which give T_e) with the S-Z data⁴. If we assume that the physical size of the cluster along the line of sight is the same as its size in the plane of the sky then, knowing the cluster redshift and angular size, we can find a value for the Hubble constant, H_0 , that is independent of the usual distance-ladder arguments^{5,6}.

The only successful measurements of the S-Z effect to date have been made with single-dish radio telescopes^{3,7,8} in which the telescope beam is switched between the direction of the cluster and a reference field. Such single-dish S-Z observations have been used to build up a one-dimensional profile of the temperature decrement through the cluster. These observations are hampered by systematic uncertainties, the most significant of which are confusion by radio sources close to the direction of the cluster, fluctuations in atmospheric emission, and uncertainty in the zero level (the signal level corresponding to the undisturbed CMBR).

Using an interferometer instead of a single-dish telescope reduces these problems significantly. For example, emission from the atmosphere is important only in that it contributes to the total noise power entering the antennas, since the emission is

uncorrelated over baselines longer than a few metres. Because an interferometer does not measure the constant background level, no spurious variation in the background level is possible in the interferometer data. Also, a suitable interferometric array can simultaneously detect both the S-Z signal (of angular size ~ 1 arcmin for a cluster at $z \sim 0.2$) with its short baselines, and the relatively compact radio sources (which will inevitably be present) with its longer baselines. Thus the contamination of the S-Z signal by confusing radio sources can be measured and removed. The array must, however, have baselines short enough to provide sufficient sensitivity to the S-Z effect; it is the lack of these short baselines that has prevented successful S-Z observations in the past with, for example, the Very Large Array⁹.

We have adapted the Cambridge 5-km telescope¹⁰ to make it suitable for observations of the S-Z effect. The upgraded instrument¹¹ has been renamed the Ryle Telescope (RT). It is an east-west synthesis telescope with eight 13-m-diameter dishes equipped with cryogenically cooled receivers operating at 5 GHz and 15 GHz, and a correlator with a bandwidth of 350-MHz. In this experiment, five of the antennas were used at an observing frequency of 15 GHz in two different configurations, each with a minimum baseline of 18 m.

We initially observed the cluster Abell 2218 with the RT between 4 December and 14 December 1992. Abell 2218 is a rich cluster at $z = 0.171$ (ref. 12), and has been observed extensively in both the X-ray and the radio portions of the spectrum. The X-ray angular diameter is ~ 2 arcmin; we expect the S-Z angular size to be similar. At 15 GHz this angular size corresponds to a baseline of ~ 20 m. The expected S-Z signal falls off rapidly with increasing baseline; we therefore configured the RT to give the largest number of its shortest (18-m) baselines, in order to provide maximum sensitivity to the S-Z effect. We did not expect to detect the S-Z effect on baselines longer than ~ 36 m; we knew from previous surveys¹³, however, that there are three significant radio sources in the field of Abell 2218, each having a flux density of a few mJy at 15 GHz and angular sizes of a few arcseconds

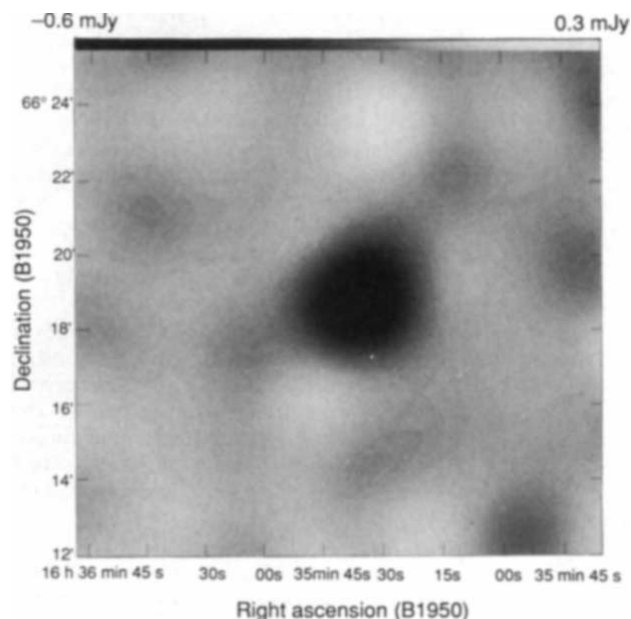
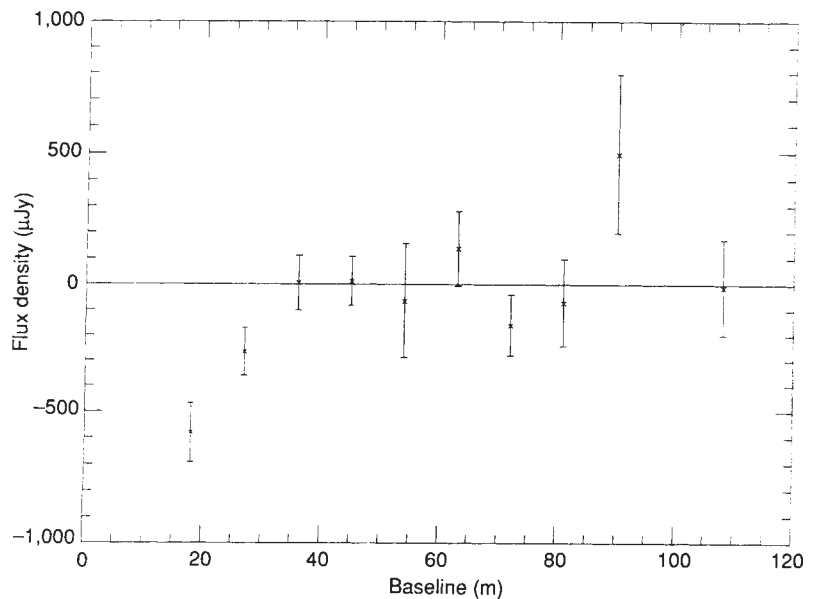


FIG. 2 A map made from the same data set as Fig. 1, but using only the 18-m baselines, and with three point-source models removed from the visibilities. Their positions and fluxes, measured only from the baselines of 36 m and longer, are: 1.83 mJy at 16 h 36 min 03.0 s, +66° 20' 26"; 0.91 mJy at 16 h 35 min 35.0 s, +66° 20' 46"; and 1.69 mJy at 16 h 35 min 08.0 s, +66° 19' 22". The map has been CLEANed and restored with a beam of 129×120 arcsec, PA = -5°. The grey-scale at the top of the figure runs from -0.6 to +0.3 mJy.

FIG. 3 The real part of the observed visibilities for each baseline, averaged over the whole data set, centred on the position of the maximum decrement of Fig. 2. These data have been calibrated and the three point-sources removed, and show a significant feature on the short baselines.



or less. (There is also a diffuse, steep-spectrum source at the cluster centre, but its flux density at 15 GHz is $<100 \mu\text{Jy}$.) A map was therefore made using the data from just the second-shortest (36-m) and longer baselines to obtain fluxes for these sources that would not be biased by the S-Z signal. The instrumental point-spread function was deconvolved from this map using the CLEAN algorithm¹⁴ and the fluxes derived from this map were then subtracted from the raw visibility data (that is, the interferometric amplitudes and phases.) A map was then made from just the 18-m-baseline data; this showed a significant negative feature close to the X-ray centre. No signal was apparent on a map made from just the 36-m and longer baselines.

We then moved the telescope into a different array to provide a baseline between 18 and 36 m. Twenty further observations were made between 16 February and 25 March 1993, and the data were reduced in a similar way. Figure 1 shows the map made from the complete data set using the longer baselines only: the three sources are clearly visible. On the 18-m baseline map (Fig. 2) a highly significant feature of $-580 \pm 110 \mu\text{Jy}$ was found centred at 16 h 35 min 35 s, $+66^\circ 18' 50''$ (B1950.0). The position of the peak decrement agrees with the centroid of the Rosat X-ray image to within the positional errors of the radio data. Using the 27-m baseline data, the level of the map at the above position was $-260 \pm 95 \mu\text{Jy}$. At the equivalent point on the 36-m-baseline map, the level was $+4 \pm 105 \mu\text{Jy}$, and on all longer baselines the levels were consistent with zero (see Fig. 3).

The combination of significant signals on two baselines, the coincidence of the feature on the map with the X-ray position, and the absence of any other such features in the field provides unambiguous evidence for the existence of an extended decrement in the CMBR towards Abell 2218. We note that this confirmation of the existence of the S-Z effect adds weight to the already strong conclusion that the CMBR originates at cosmological distances (in this case at $z > 0.171$).

We can use our data to constrain the gas distribution in Abell 2218 as follows. If we assume that the cluster gas can be described by a spherically symmetric, isothermal ' β '-model¹⁵, then the projected S-Z temperature profile as a function of angle θ on the sky will be

$$\Delta T(\theta) = \Delta T_0 \left(1 + \frac{\theta^2}{\theta_{\text{cx}}^2} \right)^{1/2 - 3\beta/2}$$

where θ_{cx} is the X-ray core radius, ΔT_0 is the temperature decrement at the centre of the cluster, and β is the ratio of the kinetic energy per unit mass in the galaxies to that in the gas (which

we treat simply as a free parameter). Fitting the visibilities predicted by such a model to our data, we find that there is a region of acceptable χ^2 (that is, χ^2 within 1 of its minimum value) in the parameter space extending from $\beta \approx 0.6$, $\theta_{\text{cx}} \approx 0.9$ arcmin, $\Delta T_0 \approx 1.1$ mK, to $\beta \approx 1.5$, $\theta_{\text{cx}} \approx 2.0$ arcmin, $\Delta T_0 \approx 0.6$ mK, the random errors being much smaller than the errors owing to the model dependence. Without further information on the shape of the S-Z profile we cannot distinguish between these possibilities. We can, however, deduce some constraints on the value of H_0 from these data. Birkinshaw and Hughes¹⁶ find $H_0 = 60 \pm 15 \text{ km s}^{-1} \text{ Mpc}^{-1}$ from observations of Abell 2218, using X-ray flux and angular size from the imaging proportional counter of the Einstein satellite, the X-ray temperature measured by the Ginga satellite¹⁷ and $\Delta T_0 = 0.6$ mK, obtained from single-dish radio observations. Simply scaling this result to fit our radio data would imply that $20 < H_0 < 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$. But applying constraints from the Einstein X-ray image to the values of the shape parameters β and θ_{cx} tends to favour the lower end of this range; to allow the upper end of the range or higher requires a profile inconsistent with the X-ray image. Given the uncertainties, however, we restrict our conclusion on H_0 to the statement that our data support a value of H_0 of 50 rather than $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The largest uncertainty arises from the assumption that the line-of-sight depth of the cluster is equal to its size in the plane of the sky. This problem can be avoided by studying a suitable randomly selected sample of clusters. Consequently, further RT observations of this and other clusters, and comparison with data from Rosat and the new ASCA satellite (Advanced Satellite for Cosmological Astrophysics; Institute for Space and Astronomical Sciences, Tokyo), will allow a better estimation of H_0 . $\square$

Received 11 May; accepted 30 July 1993.

1. Sunyaev, R. A. & Zel'dovich, Ya. B. *Comm. Astrophys. Sp. Phys.* **4**, 173–178 (1972).
2. Birkinshaw, M., Gull, S. F. & Hardebeck, H. *Nature* **309**, 34–35 (1984).
3. Birkinshaw, M. in *The Cosmic Microwave Background: 25 Years Later* (eds Mandolesi, N. & Vittorio, N.) 77–94 (Kluwer, Dordrecht, 1990).
4. Gunn, J. E. in *Observational Cosmology* (eds Maeder, A., Martinet, L. & Tamman, G.) 1–124 Geneva Observatory, Sauverny, 1978).
5. Sandage, A. & Tamman, G. A. *Astrophys. J.* **305**, 1–12 (1990).
6. de Vaucouleurs, G. in *Tenth Texas Symposium on Relativistic Astrophysics* (eds Ramaty, R. & Jones, F. C.) *Ann. N.Y. Acad. Sci.* **375**, 90–122 (1981).
7. Uson, J. M. in *Continuum Processes in Clusters of Galaxies* (eds O'Dea, C. P. & Uson, J. M.) 255–260 (National Radio Astronomy Observatory, Green Bank, 1986).
8. Herbig, T., Readhead, A. C. S. & Lawrence, C. R. *Bull. Am. astr. Soc.* **24**, 1263 (1992).
9. Partridge, R. B., Perley, R. A., Mandolesi, N. & Delpino, F. *Astrophys. J.* **317**, 112–120 (1987).
10. Ryle, M. *Nature* **239**, 435–438 (1972).
11. Jones, M. E. in *Radio Interferometry: Theory Techniques and Applications* (eds Cornwell, T. J. & Perley, R. A.) 395–399 (Astr. Soc. Pac., San Francisco, 1990).

12. Struble, M. F. & Rood, H. J. *Astrophys. J. Suppl. Ser.* **63**, 543–553 (1989).
13. Moffet, A. T. & Birkinshaw, M. *Astr. J.* **98**, 1148–1174 (1989).
14. Hogbom, J. *Astr. Astrophys. Suppl. Ser.* **15**, 417–426 (1974).
15. Cavaliere, A. & Fusco-Femiano, R. *Astr. Astrophys.* **49**, 137–144 (1976).
16. Birkinshaw, M. & Hughes, J. P. *Astrophys. J.* (submitted).
17. McHardy, I. et al. *Mon. Not. R. astr. Soc.* **242**, 215–220 (1990).

ACKNOWLEDGEMENTS. We thank the members of the Mullard Radio Astronomy Observatory, especially those who have been involved with the Ryle Telescope. We also thank the SERC for providing a substantial fraction of the funds for the RT. K.G., S.H. and D.L. acknowledge the support of SERC studentships.

Superconductivity above 150 K in $\text{HgBa}_2\text{Ca}_2\text{Cu}_3\text{O}_{8+\delta}$ at high pressures

C. W. Chu, L. Gao, F. Chen, Z. J. Huang, R. L. Meng & Y. Y. Xue

Department of Physics and Texas Center for Superconductivity, University of Houston, Houston, Texas 77204-5932, USA

RECENTLY, superconductivity at >130 K was reported^{1,2} in multi-phase samples of the compound system $\text{HgBa}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{2n+2+\delta}$ ($\text{Hg-12}(n-1)n$) with $n=1, 2, 3, \dots$. Single-phase Hg-1223 was subsequently synthesized and found to be superconducting at a record temperature of 135 K (ref. 3). It remains to be seen if a much higher transition temperature (T_c) than 135 K can be achieved in this compound system. The application of pressure generally increases the T_c of underdoped layered cuprates, with a pressure derivative of T_c that decreases with increasing doping^{4,5}. Preliminary studies on Hg-1223 ⁶ showed a pressure-induced increase in T_c without any sign of saturation up to 17 kbar. Here we report the observation of superconductivity at up to 153 K in Hg-1223 at 150 kbar, with a main transition at 147 K. This observation provides an indication that superconductivity at >150 K may be possible in this system at ambient pressure, if suitable forms of chemical substitution can be found.

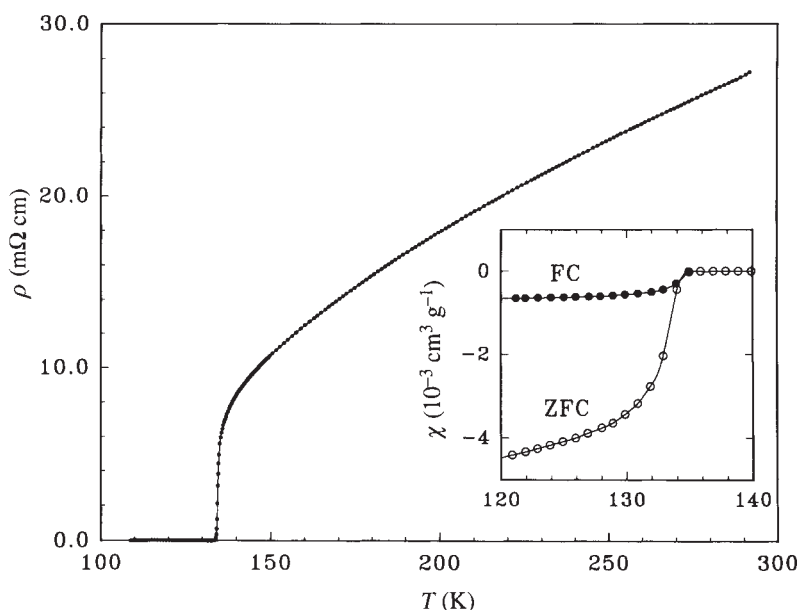
Samples were prepared, using the controlled vapour–solid reaction technique⁷ by heating a precursor pellet of $\text{Ba}_2\text{Ca}_2\text{Cu}_3\text{O}_x$ and a composite Hg source together in an evacuated quartz tube. The precursor pellets were prepared by heating and drying an aqueous solution of appropriate amounts of

$\text{Ba}(\text{NO}_3)_2$, $\text{Ca}(\text{NO}_3)_2 \cdot 4\text{H}_2\text{O}$ and $\text{Cu}(\text{NO}_3)_2 \cdot 3\text{H}_2\text{O}$ up to 620 °C in a beaker. The resulting black mixture was then ground, compacted and sintered in flowing oxygen at 900 °C for 24 to 48 hours in an alumina crucible. These precursor pellets were either used immediately or stored in a desiccator. Some of the pellets were pulverized, thoroughly mixed with HgO and compacted (in a glove bag) to form the composite Hg source (pre-reacted $\text{HgBa}_2\text{Ca}_2\text{Cu}_3\text{O}_x$). A small precursor pellet and a large composite Hg source (in a ratio of $\sim 1/3$) were sealed in an evacuated quartz tube, heated to 850 °C and kept at this temperature for 5 hours before cooling to room temperature. The Hg-1223 pellets were finally heated at 300 °C in flowing oxygen for 10 hours. X-ray diffraction revealed that $\sim 90\%$ of a pellet was Hg-1223 and the remaining $\sim 10\%$ was a mixture of CaHgO_2 , $\text{Ba}_2\text{Cu}_3\text{O}_{5.9}$ and BaCuO_2 . Both orthorhombic and tetragonal Hg-1223 , with a similar T_c , were obtained. The d.c. magnetic susceptibility (χ) and the resistivity (ρ) showed a sharp transition at ~ 135 K, with zero- ρ at 134 K (Fig. 1), corresponding to a volume fraction between $\sim 12\%$ (field-cooled, FC) and $\sim 75\%$ (zero-field-cooled, ZFC) at ~ 50 K.

Samples of $\sim 4 \times 0.8 \times 0.7$ mm were cut from a well-characterized pellet for measurements up to ~ 17 kbar under hydrostatic pressure (HP), other samples, ~ 0.9 mm diameter $\times 0.08$ mm thick were prepared for measurements up to ~ 150 kbar under quasi-hydrostatic pressure (QHP). The HP was generated inside a Teflon cup housed in a Be–Cu clamp⁸ using 3M Fluorinert as the pressure medium. The pressure was determined by a superconducting Pb manometer situated next to the sample. The QHP environment was provided by a pair of tungsten carbide anvils locked in a Be–Cu clamp⁹ using pyrophyllite as a gasket and two steatite disks as a pressure medium. The pressure was calibrated against the various phase transitions of bismuth at room temperature and the superconducting Pb manometer¹⁰ at low temperatures in separate calibration runs, and was not measured *in situ* during the experimental runs. The overall uncertainty in QHP is estimated to be $\pm 15\%$. The pressure spread across the tungsten carbide anvil surface was previously determined in our laboratory to be $\sim 10\%$ of the total pressure. A smaller sample senses a smaller spread. The superconducting transition of Hg-1223 under pressure was determined electrically by the four-probe technique. The temperature was measured by a germanium thermometer below 30 K and a chromel–alumel thermocouple above 20 K, with an uncertainty of ± 0.1 K.

Figure 2 shows the temperature dependence of the resistivity at several pressures. The samples have a slightly lower T_c than

FIG. 1 Temperature dependence of the resistivity (ρ) and d.c. magnetic susceptibility (χ) of Hg-1223 at ambient pressure.



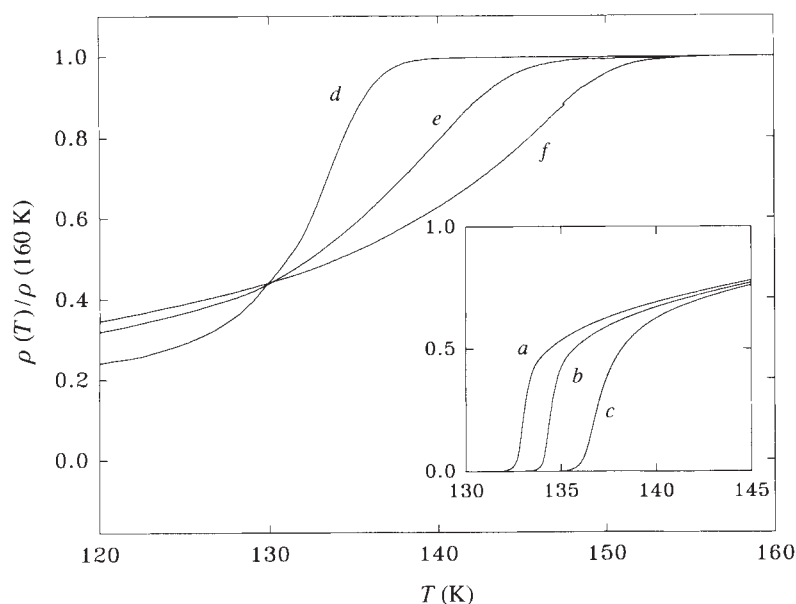


FIG. 2 Temperature dependence of the normalized resistivity of Hg-1223 under hydrostatic pressure (a, b and c correspond to 0, 5.6 and 17.1 kbar, respectively) and quasi-hydrostatic pressure (d, e and f correspond to 10, 50 and 150 kbar, respectively).

that shown in Fig. 1, perhaps owing to the nonuniformity of the pellet, or defects introduced when the samples were cut. Under HP, the sharp superconducting transition is shifted upwards in parallel as shown in Fig. 2 (curves a–c). Under QHP, when good electrical contacts were established, the superconducting transition is clearly evident but is broadened as shown in Fig. 2 (curves d–f). In addition, ρ is larger and shows a long non-zero tail below the transition. This is ascribed to possible defects introduced during thinning of the sample to ~ 0.08 mm by sand paper, and to possible microcracks generated in the sample by its setting with respect to the solid pressure medium under the QHP condition. The broad transition caused the pressure inhomogeneity across the anvil surfaces makes it difficult to define the superconducting transition temperature (T_c). Meaningful data can, however, be obtained by examining the temperature dependence of $d\rho/dT$. The onset T_c (T_{co}) and the mid-point T_c (T_{cm}) are defined in Fig. 3. To confirm that the T_{co} so defined represents the onset of a genuine superconducting transition, we have examined the transition under different measuring currents. We found that the $\rho(T)$ curves obtained with different measur-

ing currents start to diverge at T_{co} , with the high-current one shifted downward as expected for a superconducting onset, as shown in the inset of Fig. 3 for $\rho(T)$ under 150 kbar.

The pressure effect on T_{co} and T_{cm} of Hg-1223 under HP and QHP is shown in Fig. 4. T_{cm} increases reversibly with HP up to 17 kbar at a rate of ~ 0.18 K kbar $^{-1}$. Under QHP, both T_{co} and T_{cm} are enhanced initially at a rate similar to that under the HP condition. However, above ~ 20 kbar, the rate of T_c -increase reduces with pressure and becomes constant at ~ 0.02 K kbar $^{-1}$ above 80 kbar. At ~ 150 kbar, T_{co} is ~ 153 K and T_{cm} is ~ 147 K. The electrical contacts to the sample were lost due to gasket explosion upon reduction of QHP. The reversibility of the quasi-HP-effect was therefore not established.

Our results show that superconductivity can exist up to 153 K in Hg-1223 at ~ 150 kbar under quasi-hydrostatic pressure. The $T_c(P)$ behaviour of Hg-1223 is similar to that previously observed in other layered cuprates 4,5 , for which T_c increases with hydrostatic pressure and passes through a broad peak. This has been attributed to a pressure-induced change in the carrier concentration 4,5,6 . Recently, it was suggested 11 that an additional

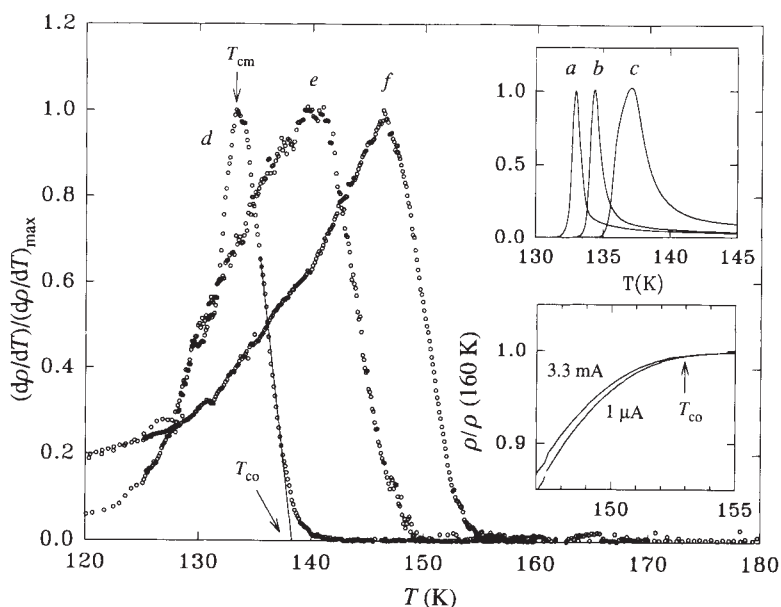


FIG. 3 Temperature dependence of normalized $d\rho/dT$ at several pressures (a, 0 kbar; b, 5.6 kbar; c, 17.1 kbar; d, 10 kbar; e, 50 kbar; f, 150 kbar) with definition of T_{co} and T_{cm} . The inset shows the normalized $\rho(T)$ for different measuring currents.

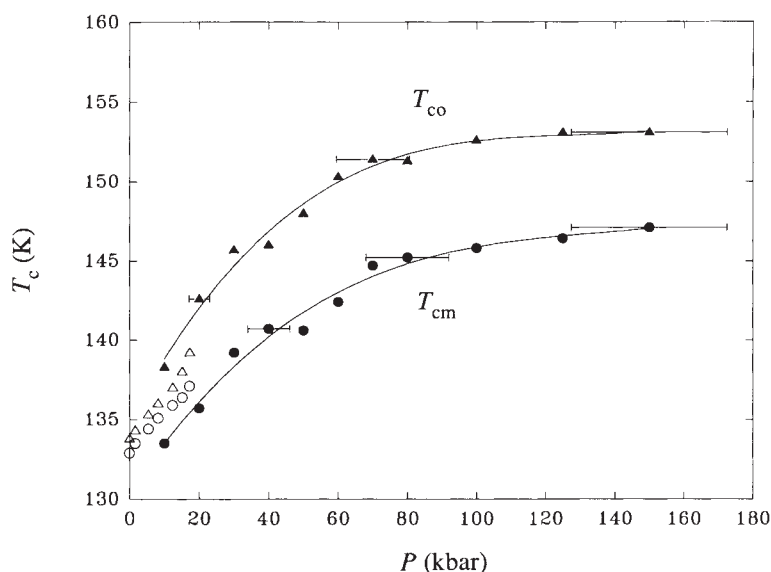


FIG. 4 Pressure dependence of T_{co} (triangles) and T_{cm} (circles) of Hg-1223 under hydrostatic (open symbols) and quasi-hydrostatic (solid symbols) pressures. The horizontal bars represent uncertainties in pressure.

explanation is needed to account for the observed pressure effect. However, a recent high-pressure study¹¹ on cation-doped $\text{YBa}_2\text{Cu}_3\text{O}_7$ suggests that the pressure-induced change in carrier concentration can account for only part of the observed pressure effect. The cause of the additional effect remains unknown. It remains to be seen whether the reduced rate of T_c enhancement above ~ 80 kbar observed here is due to T_c -suppressing defects, generated by the non-hydrostatic nature of the pressure. If this is true, T_c s above 153 K might be obtained in Hg-1223 at high hydrostatic pressures. Such a pressure effect on the T_c of Hg-1223 might also be duplicated by chemical means at ambient pressure. Such 'chemical pressure' might be applied by substituting other elements for Hg, such as those with linear oxygen coordination and/or a valence ≤ 3 . □

Received 18 August; accepted 3 September 1993.

- Schilling, A., Cantoni, M., Guo, J. D. & Ott, H. R. *Nature* **363**, 56–58 (1993).
- Gao, L. et al. *Physica C* **213**, 261–265 (1993).
- Huang, Z. J. et al. *Physica C* (in the press).
- Wijngaarden, R. J., Scholtz, J. J., van Eenige, E. N. & Griessen, R. in *Frontiers of High Pressure Research* (eds Hochheimer, H. D. & Etters, R. D.) 339–417 (Plenum, New York, 1991).
- Schilling, J. S. & Klotz, S. in *Physical Properties of High Temperature Superconductors*, Vol. 3 (ed. Ginzberg, D. M.) 59–157 (World Scientific, Singapore, 1992).
- Chu, C. W. *J. Superconductivity* (in the press).
- Meng, R. L. et al. *Physica C* (in the press).
- Chu, C. W. & Testardi, L. *J. Phys. Rev. Lett.* **32**, 766–769 (1974).
- Eichler, A. & Wittig, J. *Z. angew. Phys.* **25**, 319–327 (1968).
- Bireckoven, B. & Wittig, J. *J. Phys. E* **21**, 841–848 (1988).
- Neumeir, J. J. & Zimmermann, H. A. *Phys. Rev.* **B47**, 8386–8388 (1993).

ACKNOWLEDGEMENTS. This work was supported, in part, by the US National Science Foundation, USAFOSR, ARPA, the State of Texas and the T.L.L. Temple Foundation.

Determination of molecular symmetry in crystalline naphthalene using solid-state NMR

Julio C. Facelli*† & David M. Grant*

* Department of Chemistry and † Utah Supercomputing Institute, University of Utah, Salt Lake City, Utah 84112, USA

DIFFRACTION techniques have shown that the crystal structure of naphthalene has a unit cell with C_i symmetry^{1–7}. These studies were unable, however, to resolve any departure of the molecular structure from the D_{2h} symmetry observed in the gaseous state. We found recently⁸ that the solid-state ^{13}C -nuclear magnetic resonance (NMR) chemical shifts for naphthalene exhibit the C_i symmetry of the unit cell. If these chemical-shift data reflect structural distortions of the molecule, rather than simply intermolecular effects on the shifts owing to the C_i symmetry of the environment of each molecule, one could assert that the NMR data are able to reveal structural information beyond the limits of the diffraction methods. Here we show that this is the case by performing *ab initio* quantum-mechanical calculations of the ^{13}C chemical shifts for naphthalene, and their derivatives, with respect to structural parameters. We find that intermolecular shift terms (which of necessity exhibit C_i symmetry) can account for only about 30% of the maximum

deviations from D_{2h} symmetry; the remainder must therefore result from structural distortions of the molecules below D_{2h} symmetry. This sensitivity of NMR chemical shifts to very small changes in molecular structure opens up the possibility of using solid-state NMR along with quantum-chemical methods to refine structural parameters obtained from diffraction methods.

Naphthalene has been studied extensively by both X-ray and neutron diffraction^{1–7}, and the crystal belongs to the $P2_1/a$ space group in which the unit cell, containing two molecules, possesses C_i symmetry. Although crystal forces in this system have the potential for distorting the naphthalene structure, diffraction studies^{5,6} have been unable to measure such deformations beyond experimental errors, and the symmetry in the measured structure of this molecule is indistinguishable from the D_{2h} symmetry of gaseous naphthalene. We have recently shown⁸, however, that the ^{13}C chemical shift tensors of naphthalene measured by single-crystal NMR possess the C_i environmental symmetry of the naphthalene unit cell.

The experimental and theoretical shielding values of naphthalene, properly converted to shift values,⁹ are given in Table 1. The measured ^{13}C tensor components (δ) are reported in the icosahedral representation^{10,11}. Figure 1 defines the three-dimensional orientations of the icosahedral directions relative to the molecular frame of naphthalene and the corresponding mirror planes that would obtain in the free molecule. The wave functions, used in the theoretical calculations of the shift components, were obtained from the D_{2h} diffraction structure of naphthalene⁵, and so these theoretical shifts also exhibit D_{2h} symmetry. Some deviations from D_{2h} symmetry of individual

TABLE 1 ^{13}C chemical shift tensors in naphthalene

Experimental		Theory			
C_1	Shift value	Shift value	C_4	Shift value	$C_{1, C_4}^\dagger$ $\Delta\delta^*$
δ_1	151.04	155.97	δ_2	141.3	δ_1 2.43
δ_2	178.77	180.65	δ_1	178.6	δ_2 4.33
δ_3	167.31	166.19	δ_4	171.6	δ_3 2.16
δ_4	166.23	165.30	δ_3	171.6	δ_4 2.16
δ_5	56.43	56.40	δ_5	44.5	δ_5 3.02
δ_6	55.71	54.97	δ_6	44.5	δ_6 3.02
C_2			C_3		
δ_1	151.81	150.41	δ_2	138.6	$C_{2, C_3}^\dagger$ δ_1 2.67
δ_2	226.93	226.97	δ_1	231.1	δ_2 4.50
δ_3	125.34	125.70	δ_4	125.2	δ_3 3.41
δ_4	126.77	125.88	δ_3	125.2	δ_4 3.41
δ_5	61.71	60.59	δ_5	51.8	δ_5 2.64
δ_6	63.14	62.74	δ_6	51.8	δ_6 2.64
C_{4a}					C_{4a}
δ_1	205.66			207.7	δ_1 2.28
δ_2	207.63			207.7	δ_2 2.28
δ_3	144.80			140.8	δ_3 2.99
δ_4	144.80			140.8	δ_4 2.99
δ_5	54.11			42.2	δ_5 1.64
δ_6	52.50			42.2	δ_6 1.64

Values are given in p.p.m. in the icosahedral representation^{10,11}. The coordinate system is shown in Fig. 1. Calculated shielding values were converted to the TMS chemical-shift scale by referencing them to 191.7 p.p.m. (our 198.7 p.p.m. calculated methane value minus the 7.0 p.p.m. observed by Jameson and Jameson⁹ for the methane shift relative to TMS).

* $\Delta\delta$ values were calculated from equation (1) using two standard marginal deviations as the errors in the structural parameters⁵.

† The calculated chemical shifts for all α -, β - and bridgehead carbons, C_b , can be obtained by symmetry from those of C_1 , C_2 and C_{4a} , respectively. These symmetry considerations are explicitly included in the labelling of δ_i for the various carbons given in the table.

^{13}C chemical shift tensor elements exceed by an order of magnitude the 0.54 parts per million (p.p.m.) experimental errors⁸ in the shift measurements. The data in Table 1 demonstrate the symmetry reduction in the experimental values relative to the theoretical D_{2h} shifts. Note the differences in experimental shifts between structurally similar carbons; for example, the C_1 and C_4 tensors differ from each other as do also the related C_2 and C_3 tensors. This pronounced break from the σ_{xz} reflection plane, which would be present in D_{2h} symmetry, is readily seen in Table 1. The D_{2h} symmetry also requires that, $\delta_1 = \delta_2$ and $\delta_5 = \delta_6$ at C_{4a} owing to the σ_{yz} reflection plane passing through the bridgehead carbons. Finally, a σ_{xy} molecular reflection plane would require that $\delta_3 = \delta_4$ and $\delta_5 = \delta_6$, a condition also unobserved in the experimental values. Only $\delta_3(C_{4a}) = \delta_4(C_{4a})$, suggesting that molecular distortions from a planar structure are modest at the bridgeheads.

Two explanations can account for the observation that the ^{13}C chemical-shift data have C_i symmetry whereas D_{2h} symmetry obtains in the diffraction bond distances and angles in naphthalene. (1) Intermolecular effects on the chemical shifts, which of necessity exhibit the C_i environmental symmetry, contribute to the deviation of the shift tensor from D_{2h} symmetry. (2) The experimental C_i chemical-shift symmetry arises from structural distortions of the molecules. For the second explanation to be compatible with the diffraction data, the required structural distortions would have to be smaller than uncertainties in the measured diffraction distances in naphthalene.

Shift calculations in nonpolar and conformationally immobile molecules show that intermolecular effects usually have a magnitude of less than 1.5 p.p.m. when the intermolecular pair of interacting atoms are separated by 2.4 Å, the closest intermolecular distance in crystalline naphthalene. For example, the inter-

molecular effects on the shift tensor components produced by a C-H bond, using a CH_4 molecule positioned with the structure of the nearest adjacent C-H bond in the naphthalene unit cell (see Fig. 3 or ref. 4), affects the icosahedral shifts by 0.92 p.p.m. or less. The modest changes in shifts with a phase change from liquid to solid, which would also reflect changes in intermolecular interactions, also support this argument of small intermolecular shift interactions. Thus, while intermolecular shift terms may account for a small portion of the symmetry reduction, the largest (5 p.p.m.) variation between mirror-reflected tensor values requires structural deformations in the aromatic rings.

Molecular wave-mechanical calculations, used to predict chemical shift tensors, are very sensitive to minor changes in atomic positions. This raises the question of whether the uncertainties in atomic positions determined by diffraction are large enough to account for the shift deviations from D_{2h} symmetry. The effect of minor structural deformations on the chemical shift may be estimated by calculating the first derivative of the tensor shifts with respect to the molecular structural parameters allowed by the C_i symmetry, and then by applying standard methods for estimating the propagation of the associated structural uncertainties.

The GIAO method^{12,13} with a D95** basis set¹⁴ was used to calculate the tensor shifts and their derivatives. The use of increments in bond distances, bond angles and out-of-plane deformations, with a finite difference method, was sufficient to obtain a satisfactory approximation of the derivatives. The D_{2h} naphthalene structural parameters and their associated uncertainties, reported in the neutron diffraction data⁵, were used in this calculation. The shift derivatives, $(\partial\delta/\partial p_i)$, are similar to those previously obtained¹⁵ in smaller molecules. The cumulative uncertainties in shifts, $\Delta\delta$, were estimated as

$$\Delta\delta = \sqrt{\sum_i \left(\frac{\partial\delta}{\partial p_i} \right)^2 \Delta p_i^2} \quad (1)$$

where the Δp_i are the errors reported in the structural parameters obtained from the diffraction data. These $\Delta\delta$ values are also given in Table 1.

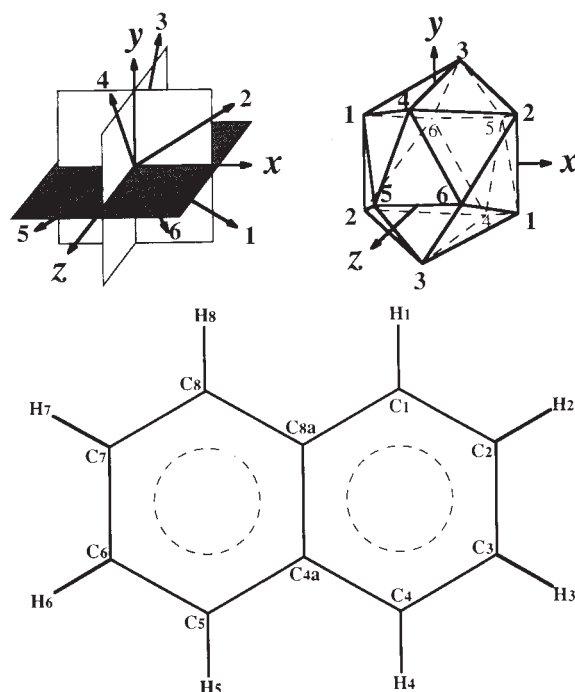


FIG. 1 The orientation of icosahedral coordinates for naphthalene, which is in the x-y plane. The numbering scheme for the carbon and hydrogen atoms is indicated.

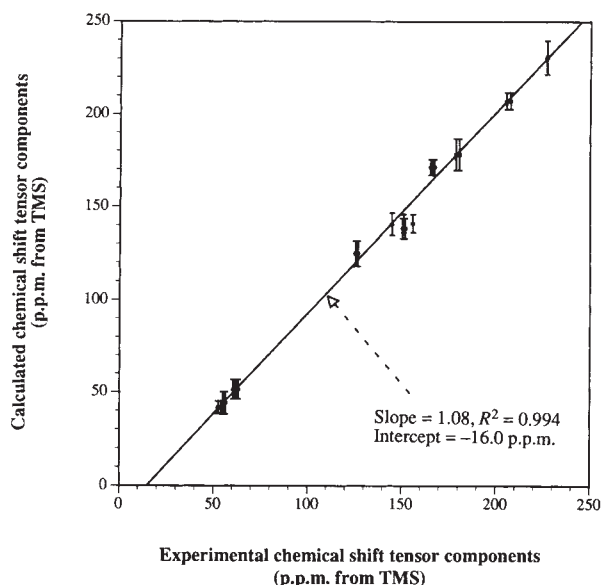


FIG. 2 The excellent correlation between experimental and calculated naphthalene ^{13}C chemical shift tensor components is shown for data in the icosahedral representation. The calculated shielding values were converted to the TMS chemical-shift scale as described in Table 1 legend. The errors in the calculated shifts were estimated from equation (1) as described in Table 1 legend. Shifts far from the bare nucleus can exhibit large absolute discrepancies between experimental and theoretical shifts. However, the 8% discrepancy, due to the neglect of electron correlation, introduces only a 0.4 p.p.m. maximum uncertainty in an incremental shifts ($\Delta\delta$) of 5 p.p.m. Such a value is comparable to the solid-state NMR measurement error of 0.54 p.p.m.

Table 1 reveals that the departures from D_{2h} symmetry observed in the experimental shift components lie within, or near to, the various calculated $\Delta\delta$ values obtained from the uncertainties in the geometrical data. Apparently geometrical variations smaller than the reported experimental errors in diffraction parameters are sufficient to account for the reduced symmetry in the chemical shift tensors. Furthermore, these predicted shift variations are much larger than the 0.54 p.p.m. experimental error in the NMR data.

Figure 2 compares the experimental naphthalene chemical shifts with the calculated values. The correlation is especially successful when error bars in the calculation (due to uncertainties in the structural parameters) are considered. An obvious discrepancy, due to the neglect of electronic correlation effects, accounts for a $\sim 8\%$ systematic overestimate of the predicted shifts.¹⁶ This effect skews the fitted slope in Fig. 2 from unity but does not appear to degrade the overall correlation. Thus, Fig. 2 strongly supports the assertion that structural distortions can account for much of the symmetry reduction in these shift tensors, and that uncertainties in the structural parameters (used in the wave functions) become a limiting factor in the calculations of chemical shifts.

We conclude that the naphthalene molecule in the crystal unambiguously exhibits the C_i symmetry of the naphthalene unit cell, and while part of the chemical shift data can be explained by intermolecular interactions, the dominant effect on shift symmetry seems to arise from small distortions in the molecular structure that are immeasurable in the diffraction data. Theoretical estimates of the ^{13}C chemical shifts exhibit a high intrinsic dependence on the geometry of the molecule, and measured chemical shifts coupled with theory have the potential to provide structural information. Improved computational tools and theoretical methods for calculating chemical shifts in crystalline systems are rapidly advancing, and should shortly be able to predict chemical shift tensors with an accuracy under 1 p.p.m. for moderately sized molecules such as naphthalene. Improved

computational capabilities can also be expected to reduce the electron correlation problem. Such theoretical methods will then allow ^{13}C -NMR shift tensors from single crystals to be used in the quantitative refinement of molecular structures obtained by diffraction techniques.

The use of solid-state NMR methods for refining structural data should be especially attractive in the study of biomolecules with molecular weights in the range $10\text{--}20 \times 10^3$, where diffraction data have even larger structural errors. Many types of crystalline imperfections, which degrade diffraction data, have no effect on chemical-shift data that are not sensitive to such imperfections as translational disorder or the absence and/or occlusion of a given molecular impurity. A ^{13}C -labelled atom in a large molecule may be observed with the magnification factor of the isotopic enrichment, allowing one to focus on the active sites of larger molecular systems while avoiding spectral interference from less relevant parts of the molecule. $\square$

Received 17 December 1992; accepted 16 July 1993.

1. Bragg, W. H. *Proc. phys. Soc.* **34**, 33–50 (1921).
2. Robertson, J. M. *Proc. R. Soc. A* **142**, 674–688 (1933).
3. Abrahams, S. C., Robertson, J. M. & White, J. S. *Acta crystallogr.* **2**, 233–238 (1949).
4. Cruickshank, D. W. J. *Acta crystallogr.* **10**, 504–508 (1957).
5. Pawley, G. S. & Yeats, E. A. *Acta crystallogr.* **B25**, 2009–2013 (1969).
6. Ponomarev, V. I., Filipenko, O. S. & Atovmyan, L. O. *Soviet Phys. Crystallogr.* **21**, 215–216 (1976).
7. Brock, C. P. & Dunitz, J. D. *Acta crystallogr.* **B38**, 2218–2228 (1982).
8. Sherwood, M. S., Facelli, J. C., Alderman, D. W. & Grant, D. M. *J. Am. chem. Soc.* **113**, 750–753 (1991).
9. Jameson, A. K. & Jameson, C. J. *Chem. Phys. Lett.* **134**, 461–466 (1987).
10. Alderman, D. W., Sherwood, M. H. & Grant, D. M. *J. magn. Reson.* **A101**, 188–197 (1993).
11. Grant, D. M., Facelli, J. C., Alderman, D. W. & Sherwood, M. H. in *Nuclear Magnetic Shielding & Molecular Structure* (ed. Tossell, J. A.) NATO-ARW Series **C386**, 367–384 (1993).
12. Ditchfield, R. *Molec. Phys.* **27**, 789–807 (1974).
13. Wolinski, K., Hinton, J. F. & Pulay, P. *J. Am. chem. Soc.* **112**, 8251–8260 (1990).
14. Dunning, T. H. & Hay, P. J. *Methods of Electronic Structure Theory* (ed. Schaefer, H. F. III) 1–28 (Plenum, New York, 1977).
15. Chestnut, D. B. A. *Rep. NMR Spectrosc.* **21**, 51–97 (1989).
16. Orendt, A. M. et al. *J. Am. chem. Soc.* **114**, 2832–2836 (1992).

ACKNOWLEDGEMENTS. This work was supported by the NIH and the US Department of Energy. Computer resources from the Utah Supercomputing Institute were funded by the State of Utah.

Evidence for two kinds of hydrogen bond in ice

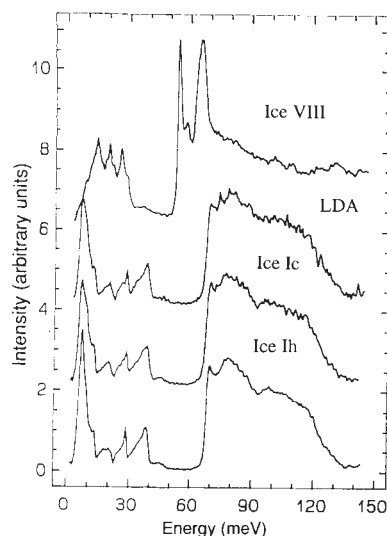
Jichen Li*† & D. K. Ross*

* Department of Pure and Applied Physics, University of Salford, Salford M6 4WT, UK

† School of Physics and Space Research, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

DESPITE its simplicity at the molecular level, water is a complex and poorly understood liquid¹. The reasons for this centre around the existence of a dynamic hydrogen-bonded network throughout the liquid. Attempts to describe the structure of liquid water have tended to invoke either continuum models such as the 'distorted-bond' model², which assumes that the hydrogen-bonded structure relaxes on a timescale similar to that in other liquids, and mixture models, such as the 'flickering-cluster' model³, which postulate the coexistence of two or more long-lived structures in the liquid. Here we analyse inelastic neutron-scattering spectra for the ice I solid phase of water, which provide evidence for the existence of two different kinds of hydrogen-bond, of different strengths, in the solid. A model in which strong and weak hydrogen-bonds in the ratio of about 2:1 are randomly distributed throughout the network is able to reproduce the neutron spectra. If we can assume that the same kind of bimodal hydrogen-bonding exists in the liquid state, our model may be able to explain several of the anomalous properties of liquid water, such as the large specific heat and the unusual behaviour of water in thin films and clusters.

FIG. 1 IINS measurements of ice Ih, Ic and LDA showing that these are almost identical in the molecular optic modes region, 20–40 meV, indicating that the two types of H-bond exist in all three structures. For ice VIII, the high-energy peak at 36 meV is missing.



Although in the past thirty years there have been a large number of *ab initio* quantum-mechanical (molecular orbital) calculations for water dimers, trimers and larger clusters^{4–6}, there has been no attempt to make these calculations for pairs of water molecules with different configurations in a realistic environment such as real lattice. Hence averaged pair potentials have been used in almost all theoretical models⁷. There has been some speculation about the possibility of strong and weak hydrogen-bonds²⁰, but without firm experimental evidence.

The vibrational spectra of the different phases of ice have frequently been investigated using infrared and Raman^{8,9} techniques, but, owing to the proton disorder in most such structures, the normal selection rules governing the interaction of radiation with these lattices are broken and so any analysis of the spectra is difficult and sometimes misleading. In contrast, inelastic incoherent neutron scattering (IINS) is a more direct probe, because the IINS spectrum is directly proportional to the phonon density of states (PDOS) weighted by the mean-square amplitude associated with each mode, and because the PDOS derived directly from the measured IINS cross-section can be rigorously calculated for a particular model of the lattice dynamics. In the past, there have been a number of IINS measurements for ice Ih^{10,11} and for water¹². There are also recent measurements of IINS spectra for ices¹³. Such measurements, however, were often of low resolution and had poor statistics, so the spectra were lacking in the detail that is crucial to any interpretation

of the translational region below 40 meV. We have been conducting systematic studies of the dynamics of various ice phases, including most of the 'exotic' high-pressure structures^{14–17}, in order to measure the interatomic force constants at different O–O separation distances and hence allow the construction of an effective pair potential. Figure 1 shows the spectra obtained for ice Ih¹⁴, Ic (this work), VIII¹⁹ and low-density amorphous (LDA) ice (this work) using the TFXA spectrometer on the pulsed neutron source ISIS at the Rutherford Appleton Laboratory. The IINS spectra of ice Ic and LDA ice show two well separated molecular-optic bands which peak at 27 and 36 meV with unusually pronounced triangular shapes (identical to the spectrum of ice Ih); rather than the single main peak at 27 meV, as is observed in the Raman and infrared spectra (Fig. 2) where the peak at 36 meV appears only as a shoulder on the right side of the main peak. Polarized spectra of single ice crystals show that the intensities of these two molecular-optic peaks are independent of the crystal orientation¹⁴ which rules out the possibility of two different force constants in the *a* and *c* directions as suggested by Renker¹¹. The spectra of other forms of proton-disordered ice, such as ice II and V, have similar features¹⁵. Conventional explanations of the shoulder at 36 meV in the Raman and infrared spectra, such as the suggestion that the shoulder is due to transverse optic and longitudinal optic splitting as is found in ionic crystals, cannot explain these results. Our attempts to reproduce these IINS spectra using lattice dynamic calculations are also far from satisfactory because they yield only one peak¹⁸. We are therefore forced to the conclusion that the existence of two molecular optic bands in the spectrum implies two strengths of H-bonds.

FIG. 3 A section of ice I lattice is used to show the distribution of the strong and the weak H-bonds. The solid lines represent a strong bond and the pair of thin lines represent a weak bond. Four possible orientations of molecule pairs are also shown: A and B are along the *c*-axis of the ice Ih structure and C and D are in the basal plane. In ice Ic, only C and D types of molecule pairs are found.

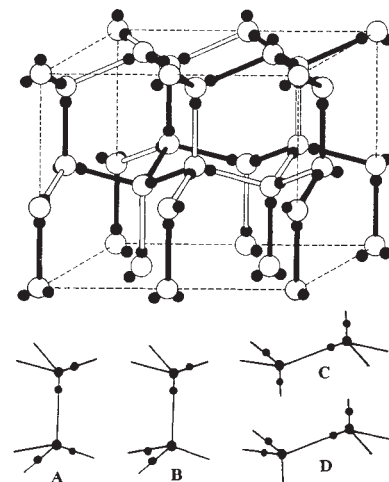
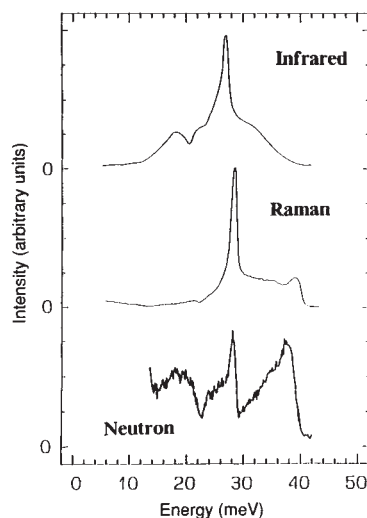


FIG. 2 Comparison of typical infrared, raman and neutron spectra for ice Ih showing that there are significant differences in emphasis in the different techniques. The infrared and raman spectra show that the main peak is at 27 meV, whereas the neutron data shows that the high-energy peak at 36 meV contains twice as much of the phonon density of states as the lower-energy peak at 27 meV.



We have made a number of lattice dynamic calculations using two H-bond force constants in a large superlattice unit cell containing 64 molecules (8 unit primary cells), where the protons are distributed in a random way, but subject to the Bernal-Fowler ice rules²⁴. The empirical model treats the H₂O molecules as point masses. There are therefore only three force constants, the O–O bond-bending force constant, $G=0.56$ eV rad⁻², and the O–O bond-stretching force constant, κ , where κ will have the values $\kappa_1=1.1$ eV Å⁻² and $\kappa_2=2.1$ eV Å⁻², which are statistically distributed through the lattice. We assume that the two force constants are related to the orientations of two adjacent molecules, described in terms of the orientation of the dipoles associated with the molecules as shown in Fig. 3. In ice Ih, there are arrangements A, B, C and D, whereas in ice Ic, only C and D are found. If the interaction were purely electrostatic, B and C would be strong interactions while A and D would be weak. Moreover, in a random lattice, these two types (weak and

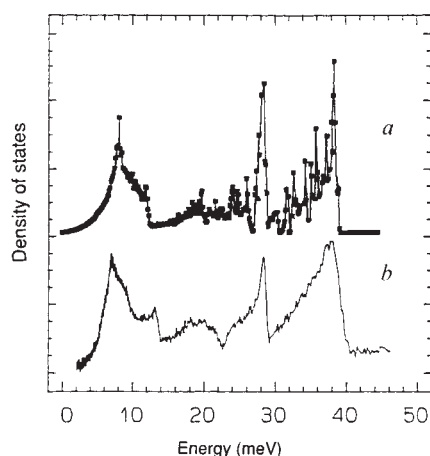


FIG. 4 Upper curve *a* is the calculated phonon density of states for ice Ih, showing two triangular peaks as observed in the measured data *b*. The use of a larger superlattice and finer mesh points in *q* will further improve the smoothness of the calculated density of states.

strong) would appear in ice in the required ratio of $\frac{1}{3}:\frac{2}{3}$. A computer program PHONON (M. Leslie, personal communication) was then used to calculate not only the dispersion curves, but also the PDOS — by integrating over all the *q* points in the first Brillouin zone. In order to reproduce the scattering correctly, the PDOS must also be weighted for each atom by the amplitude-squared for each eigenvalue at each *q* value times the scattering cross-section for that atom. The resulting PDOS (shown in Fig. 4) is in remarkable agreement with the measured inelastic scattering. The calculated spectrum shows a lot of fine structure before convolution with the resolution function of the instrument, but the use of a larger superlattice and finer mesh points in *q*-space within the first Brillouin zone would produce a smoother function. The two bands at 27 and 36 meV are well reproduced. The distinct triangular shapes both have sharp cut-offs on the high-frequency side of the bands. In addition, the ratio of the integrated peak areas are determined entirely by the ratio of strong/weak H-bonds as distributed through the superlattice cell. Because a superlattice was used, the individual dispersion curves are not easy to identify or to analyse. Hence, to give an impression of the dispersion involved in the model, we have picked out a few subsets of the superlattice corresponding to the primary unit cells and have calculated the corresponding dispersion curves¹⁶. These show flat upper energy curves at 27 and 36 meV, which will give the sharp edges of the two bands in the IINS spectrum: the integrated areas under the two peaks are proportional to the corresponding dispersion curve densities. Furthermore, for one of the ordered forms of ice, ice VIII (ref. 19), the calculated spectrum shows no high-energy peak at 36 meV because the structure, being proton-ordered, has only the dipole arrangement (D) in Fig. 3 corresponding to our weak force constant. Here, the calculation shows one peak in this

region which corresponds to the lower-energy peak at 27 meV in Fig. 2.

The only hypothesis in the above model is that there are two different strengths for the H-bonds in the structures, which are randomly distributed throughout the lattice. (We also assume that the strong force constant will involve a larger bond energy.) The large difference between the force constants of the H-bonds cannot be produced by either dipole-dipole or other electrostatic interactions, which would give a maximum difference of 15%. This implies that the use of averaged pairwise interaction potentials between hydrogen-bonded water molecules is unlikely to reproduce the properties of ice or water. In the past, the possible existence of a 'strong' bond in water has often been discussed and a number of structural models, such as the stage models²⁰, are actually based on a mixture of strong and weak H-bonds. The strengths of the H-bonds depend on how many H-bonds are actually broken in the liquid. Stanley and Teixeira²¹ have proposed a percolation model which has been considered to be one of the two-state models. Our observation cannot be regarded as supporting evidence for the two-state models as originally defined, because the difference in strength of the H-bonds occurs in solid-state ice, that is in fully connected systems. However, our proposed network of strong and weak bonds looks rather like the percolation bond of Stanley (Fig. 3), but the difference is that the bond in the percolation model is either absent or present.

We also suggest that bifunctional H-bond interaction should not be very different in the liquid state. (Experimentally, however, the IINS spectrum of water is normally featureless owing to the recoil effect of the water molecules, because the measurements are usually made at relatively high wave vector (*Q*) values, so the proposal is hard to test.) Such a hypothesis could explain a wide range of the anomalous properties of water. The high heat capacity, for example, might be explained by changes in the ratio of strong/weak bonds, which will have different bond energies. As the temperature increases, the number of weak bonds will increase, and the extra static energy stored in the weak bonds will give a high heat capacity. The large temperature range of liquid water might be explained if melting requires the breaking of the weak bonds only, whereas boiling requires the breaking of both weak and strong bonds. In thin films of water, such as those on surfaces or porous silica or on biological membranes, the reorientation of water molecules can take place relatively easily and it may be that only strong H-bond form. Our IINS spectra of water layers adsorbed in porous Vycor glass²² show only a symmetric peak at 37 meV, which indeed suggests that all H-bonds are of the strong variety. These strongly H-bonded clusters should have all the properties reported for vicinal water²³.

We believe that our empirical model provides insight into the complex nature of the H-bond in water. The fact that electrons in *sp*³ orbitals of oxygen atoms can easily be rehybridized to respond to the relative configurations of adjacent molecules may account for the two kinds of H-bond. This change in electronic states may also affect the internal O-H covalent bond strength. □

Received 28 May; accepted 5 August 1993.

1. Frank H. S. in *Water: A Comprehensive Treatise* Vol. 1 (ed. Franks, F.) 515–577 (Plenum, New York & London, 1972).
2. Pople, J. A. *Proc. R. Soc. A* **205**, 163–171 (1957).
3. Frank, H. S. & Wen, W.-Y. *Discuss. Farad. Soc.* **24**, 133–148 (1957).
4. Morokuma, K. & Pedersen, L. J. *chem. Phys.* **48**, 3275–3282 (1968).
5. Millot, C. & Stone, A. J. *Molec. Phys.* **77**, 439–462 (1992).
6. Lee, C. et al. *Phys. Rev. Lett.* **69**, 462–465 (1992).
7. Finney, J. L., Quinn, J. E. & Baum, J. O. in *Water Science Review* Vol. 1 (ed. Franks, F.) 93–170 (Cambridge Univ. Press, 1985).
8. Whalley, E. & Bertie, J. E. *J. chem. Phys.* **46**, 1271–1284 (1967).
9. Mineeva-Sukarova, B., Sherman, W. F. & Wilkinson, G. R. *J. Phys.* **C17**, 5833–5850 (1984).
10. Prask, H., Boutin, H. & Yip, S. J. *chem. Phys.* **48**, 3367–3376 (1968).
11. Renker, B. in *Physics and Chemistry of Ice* (eds Whalley, E., Hones, S. J. & Gold, L. W.) 82–89 (Univ. of Toronto Press, 1973).
12. Chen, S.-H. *Proc. of the NATO Adv. Study Inst. on Hydrogen-bonded Liquids* (eds Dore,

- J. C. & Teixeira, J.) 289–295 (Kluwer Academic, London, 1991).
13. Severson, E. C. et al. in *Phonon 89* (eds Hunklinger, S., Lugwig, W. & Weiss, G.) 537–541 (World Scientific, London, 1990).
14. Li, J.-C., Ross, D. K., Hall, P. G. & Tomkinson, J. J. *phys. B* **156/157**, 376–379 (1989).
15. Li, J.-C. et al. *J. chem. Phys.* **94**, 6770–6775 (1991).
16. Li, J.-C. & Ross, D. K. *Int. Conf. on Phys. and Chem. of Ice* (eds Maeno, N. & Hondoh, T.) 27–34 (Hokkaido Univ. Press, 1992).
17. Li, J.-C. et al. *J. Phys. Condensed Matter* **4**, 2109–2116 (1992).
18. Marchi, M., Tse, J. S. & Klein, L. J. *chem. Phys.* **85**, 2414–2419 (1986).
19. Kolesnikov, A. I. et al. *Phys. Lett. A* **168**, L308–L312 (1992).
20. Kistenmacher, H., Popkie, H. & Clementi, E. *J. chem. Phys.* **60**, 4455–4465 (1972).
21. Stanley, H. E. & Teixeira, J. J. *chem. Phys.* **73**, 3404–3422 (1980).
22. Li, J.-C. et al. *Inst. of Phys. Conf. Ser. Vol. 101* (eds Fairbanks, M. C., North, A. N. & Newport, R. J.) 109–118 (Arrowsmith, Bristol, 1990).
23. Etzler, F. M. & Drost-Hanson, W. *Croatica Chemica Acta* **56**, 563–592 (1983).
24. Bernal, J. D. & Fowler, R. H. *J. chem. Phys.* **1**, 515–548 (1933).

Simulating the critical behaviour of complex fluids

J. Ilja Siepmann*, Sami Karaborni & Berend Smit†

Shell Research BV, Koninklijke/Shell-Laboratorium, Amsterdam, PO Box 3003, 1003 AA Amsterdam, The Netherlands

ALTHOUGH the liquid–gas phase equilibria of simple fluids have been studied extensively since the seminal work of van der Waals, the properties of fluids with more complex molecular structures, such as polymers, present a less tractable problem both theoretically and experimentally. The phase behaviour of hydrocarbons is of particular importance for the petrochemical industry. But despite significant experimental and theoretical efforts, the phase diagrams of the straight-chain alkanes longer than decane (C_{10}) are known only partially, and even qualitative aspects such as the chain-length dependence of the critical properties are poorly understood. Until recently it was considered impossible to estimate the critical properties of such complex fluids using computer simulations. Here we report Monte Carlo simulations of the phase diagrams of alkanes with unbranched carbon chains as long as C_{48} , up to the vicinity of the liquid–vapour critical points. Our calculations show that, in contrast to the traditional view, the critical density of the long-chain alkanes decreases rather than increases with carbon number. This work indicates that simulations can be used as an ‘engineering tool’ to estimate properties that are not readily accessible experimentally.

Alkanes are thermally unstable above ~ 650 K, which makes experimental determination of the critical points of alkanes longer than decane (C_{10}) extremely difficult, as these tend to lie at still higher temperatures. The longer alkanes, however, are present in mixtures of practical importance for the petrochemical industry. In these mixtures, the number of components can be so large that it is not practical to determine all phase diagrams experimentally. It is therefore necessary to rely on predictions made by equations of state. The parameters of these equations are directly related to the critical properties of the pure components. Therefore, the critical properties of the long-chain alkanes are essential in the design of petrochemical processes, even if they are unstable close to the critical point. Unfortunately, experimental data are scarce and contradictory, and one has to rely on semi-empirical methods to estimate the critical properties¹. Because we can use our simulation technique to study phase behaviour of the longer alkanes at conditions where experiments are not (yet) feasible, we are in a position to make predictions of the critical properties of these molecules.

The simulation technique used here is a combination of the Gibbs-ensemble technique^{2,3} and the configurational-bias Monte Carlo method^{4,5}, the details of which are described in refs 6 and 7. In the Gibbs-ensemble scheme, simulations of the liquid and vapour phases are carried out in parallel. Monte Carlo rules which allow for changes in the number of particles and the volume, ensure that the two boxes (liquid and vapour) are in thermodynamic equilibrium with each other. Because the two boxes are not in ‘physical contact’, there is no interface, and the bulk properties of the two coexisting phases can be obtained directly with a surprisingly small number of particles. This makes the Gibbs ensemble extremely efficient for phase-equilibrium calculations. The major limitation of the Gibbs-ensemble technique is that one of the steps involves the exchange of particles between the two boxes. For liquids consisting of small molecules this does not cause serious problems. For chain molecules, however, the probability of successful exchanges can become very small. For example, under conditions where it takes approximately 10^3 attempts per successful exchange of a methane molecule, it takes

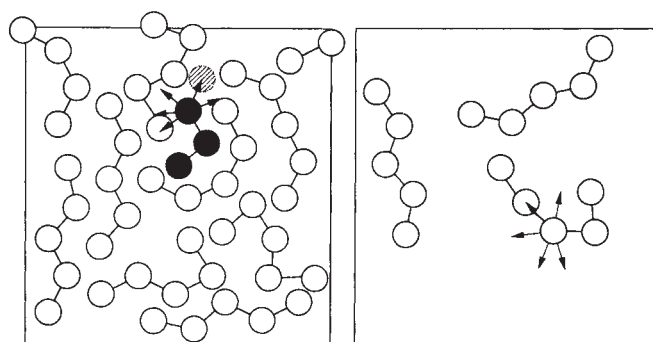


FIG. 1 Schematic drawing of the configurational-bias Monte Carlo algorithm. The figure shows an attempt to exchange a particle from the gas phase (right) into the liquid phase (left). Three segments of the (dark) chain have been grown successfully in the liquid phase and k trial positions are generated (indicated by arrows) to insert the fourth segment. For each of these positions, the energy u_i is calculated and one position is selected (in this case it will be the one with the light shading) with a probability $P_i(l) = \exp(-u_i/k_B T) / W_{\text{new}}(l)$ with $W_{\text{new}}(l) = \sum_{j=1}^k \exp(-u_j/k_B T)$, where T is the temperature and k_B is Boltzmann's constant. This is done for all segments and the factor $W_{\text{new}} = \prod_l W_{\text{new}}(l)$ is determined. For the old configuration $W_{\text{old}} = \prod_l W_{\text{old}}(l)$ is calculated by retracing the ‘path’ of the chain (see diagram of the gas phase, right). For each segment, $k-1$ trial positions are generated which together with the actual position of the segment are used to calculate $W_{\text{old}} = \sum_{j=1}^k \exp(-u_j/k_B T)$. It can be proved that the bias of the insertion can be removed by replacing $\exp(-\Delta U/k_B T)$ in the acceptance rule by $W_{\text{new}}/W_{\text{old}}$ (ref. 6).

of the order of 10^{3n} attempts for a linear alkane with n carbon atoms. The conventional scheme is based on a random insertion of a molecule. In our algorithm, the insertion of the chain is biased. A chain is ‘grown’ atom-by-atom such that regions of favourable energies are found (Fig. 1). The bias is then removed by adjusting the acceptance rules⁶. As a result, the number of successful exchanges is enhanced by an order of magnitude for pentane (C_5) and up to 15 orders of magnitude for octatetracontane (C_{48}).

The Gibbs-ensemble simulations have been performed in cycles. In each cycle a number (equal to the number of particles) of randomly selected Monte Carlo moves are attempted: displacements of a molecule, rotation of a molecule, volume change

TABLE 1 Details of the n -alkanes model

Physical aspect	Potential function	Parameters
Non-bonded interactions	$U_{\text{LJ}}(r_{ij}) = 4\epsilon_{ij}((\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^6)$	$\sigma_{\text{CH}_3} = \sigma_{\text{CH}_2} = 3.93 \text{ \AA}$ $\epsilon_{\text{CH}_3} = 114.0 \text{ K}$ $\epsilon_{\text{CH}_2} = 47.0 \text{ K}$
Bond-bending	$U_{\text{bending}}(\theta_i) = 1/2 K_{\theta}(\theta_i - \theta_{\text{eq}})^2$	$k_{\theta} = 62,500 \text{ K rad}^{-2}$ $\theta_{\text{eq}} = 114^\circ$
Angle torsion	$U_{\text{torsion}}(\phi_i) = a_1(1 + \cos \phi_i) + a_2(1 - \cos(2\phi_i)) + a_3(1 + \cos(3\phi_i))$	$a_1 = 355.03 \text{ K}$ $a_2 = -68.19 \text{ K}$ $a_3 = 791.32 \text{ K}$

The model is based on the united-atom description, that is methyl or methylene groups are considered as single interaction centres. The pseudo-atoms in a given chain are assumed to be connected by rigid bonds ($d_{\text{CC}} = 1.54 \text{ \AA}$). The non-bonded interactions are described by Lennard–Jones potentials which are truncated at $13.8 \text{ \AA}$. The usual tail corrections¹⁵ are used to estimate the interactions beyond the cut-off. The parameters for the unlike interactions have been calculated using $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$. Bond-bending is modelled by a harmonic potential¹⁶ and changes in the torsional angles are controlled by the Jorgenson potential¹¹.

* Present address: Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323, USA.

† To whom correspondence should be addressed.

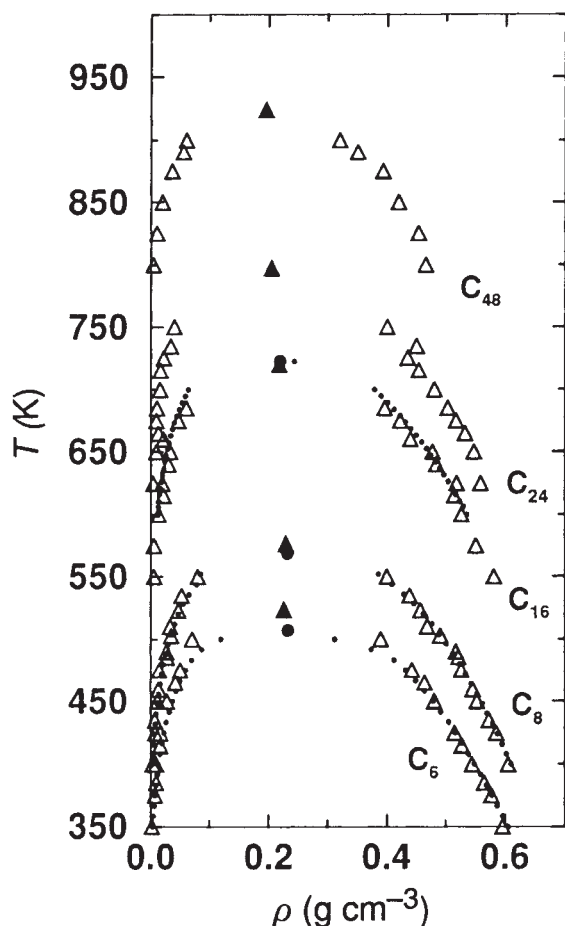


FIG. 2 The vapour-liquid curves of various alkanes, plotted as temperature T against density ρ . Simulations are represented by open triangles and the critical point by a large filled triangle. The experimental results are represented by black dots (C_6 , C_8 are from ref. 17 and the data for C_{16} are estimates from an equation of state). The experimental critical points (represented by large filled circles) are from ref. 14. The simulations were performed with numbers of molecules ranging from 200 (for C_5), to 100 (for C_{48}); the number of trial orientations in the configurational-bias Monte Carlo scheme range from 5 (for C_5) to 10 (for C_{48}). The number of equilibrium cycles was at least 3,000 for C_5 ranging to 6,000 for C_{48} , and production runs were performed over at least 5,000 cycles for all molecules.

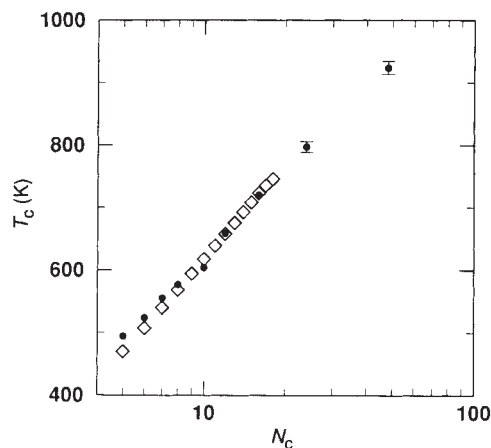


FIG. 3 Critical temperature of the normal alkanes (T_c) as a function of the carbon number (N_c). The experimental data of Anselme *et al.*¹⁴ and Steele¹³ are represented by squares and the results from the simulations by filled circles. Error bars are shown when larger than the symbol size.

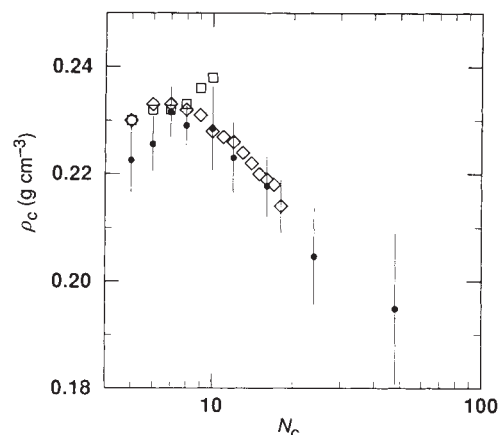


FIG. 4 Critical density of the normal alkanes (ρ_c) as a function of the carbon number (N_c). The experimental data of Anselme *et al.*¹⁴ are represented by diamonds, the data of Steele¹³ by squares and the results from the simulations by filled circles. Note that for C_5 the data of Anselme *et al.* and Steele coincide.

of the two boxes, re-growing parts of a molecule, and exchanging molecules between the two boxes. For the latter two moves configurational-bias Monte Carlo has been used. For systems with strong intramolecular interactions it is important to take these interactions into account when generating the trial orientations in the configurational-bias Monte Carlo scheme^{5,8}. The probability of acceptance of an exchange move ranged from 2–10% for C_5 to 0.5–3% for C_{48} , depending on the temperature. The results of the Gibbs-ensemble simulations are analysed using the techniques described in ref. 9 and the critical point is estimated using the procedure given in ref. 10.

Phase diagrams are very sensitive to the choice of interaction potentials. Most available models for alkanes have been obtained by fitting simulation data to experimental properties of the liquid at standard conditions. We have tested several of the published alkane models, but it turned out that none of them was able to reproduce the liquid-vapour coexistence properties for both short and long chains. For instance, the models of refs 11 and 12 give nearly identical liquid properties in standard conditions, but yield estimates of the critical temperature of octane that differ by 100 K. For this reason, we developed a new set of energy and size parameters of the intermolecular potential, the intramolecular potentials being taken directly from the literature. Details of the potentials are given in Table 1. The coexistence curves, shown in Fig. 2, demonstrate that our model is in satisfactory agreement with the experimental data.

In Fig. 3 the critical temperature is plotted against the carbon number. The simulations reproduce the experimental critical temperatures very well. There is, however, considerable disagreement between the various experimental estimates of the critical densities. Much of our current knowledge of the critical properties of the higher alkanes is based on extrapolations of fits to the experimental data up to C_8 . The most widely used extrapolations assume that the critical density is a monotonically increasing function of the carbon number, approaching a limiting value for the very long alkanes^{1,13}. In contrast to such predictions based on extrapolation, recent experimental data¹⁴ indicate that the critical density reaches a maximum at C_8 and then decreases monotonically. Other data¹³, however, do not give any evidence for such a maximum (Fig. 4). Our simulations indicate the same trend as that observed by Anselme *et al.*¹⁴. Because calculations can also be done for much longer alkanes, our results strongly support the latter experiments.

Our work shows that simple group-contribution like intermolecular potentials can be developed to model systems of

industrial importance accurately. Simulations could therefore become an important technique for estimating properties of systems that are difficult to study by traditional methods. The techniques employed here are not limited to *n*-alkanes but can be applied in many other computer simulations of complex fluids. □

Received 15 June; accepted 9 August 1993.

1. Tsouropoulos, C. *A.I.Ch.E. J.* **33**, 2080–2083 (1987).
2. Panagiotopoulos, A. Z. *Molec. Phys.* **61**, 813–826 (1987).
3. Panagiotopoulos, A. Z., Quirk, N., Stapleton, M. & Tildesley, D. J. *molec. Phys.* **63**, 527–545 (1988).
4. Siepmann, J. I. & Frenkel, D. *Molec. Phys.* **75**, 59–70 (1992).
5. Frenkel, D., Mooij, G. C. A. M. & Smit, B. J. *Phys.: Condensed Matter* **4**, 3053–3076 (1992).

6. Mooij, G. C. A. M., Frenkel, D. & Smit, B. J. *Phys.: Condensed Matter* **4**, L255–L259 (1992).
7. Laso, M., de Pablo, J. J. & Suter, U. W. J. *chem. Phys.* **97**, 2817–2819 (1992).
8. Siepmann, J. I. & McDonald, I. R. *Molec. Phys.* **79**, 457–473 (1993).
9. Smit, B., de Smedt, Ph. & Frenkel, D. *Molec. Phys.* **68**, 931–950 (1989).
10. Smit, B. & Williams, C. P. J. *Phys.: Condensed Matter* **2**, 4281–4288 (1990).
11. Jorgensen, W. L., Madura, J. D. & Swenson, C. J. *J. Am. chem. Soc.* **106**, 6638–6646 (1984).
12. Toxvaerd, S. J. *chem. Phys.* **93**, 4290–4295 (1990).
13. Tsouropoulos, C. & Tan, Z. *Fluid Phase Equilibria* **83**, 127–138 (1993).
14. Anselme, M. J., Gude, M. & Teja, A. S. *Fluid Phase Equilibria* **57**, 317–326 (1990).
15. Allen, M. P. & Tildesley, D. J. *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
16. Van der Ploeg, P. & Berendsen, H. J. C. *J. chem. Phys.* **76**, 3271–3276 (1982).
17. Smith, B. D. & Srivastava, R. *Thermodynamics Data for Pure Compounds: Hydrocarbons and Ketones* (Elsevier, Amsterdam, 1986).

ACKNOWLEDGEMENTS. We thank M. J. Anselme, D. Frenkel, W. G. Heitman, A. K. van Helden, H. P. C. E. Kuipers and S. Schreuder for comments.

The influence of water on the petrogenesis of subduction-related igneous rocks

Glenn A. Gaetani*, Timothy L. Grove* & Wilfred B. Bryan†

* Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Woods Hole Oceanographic Institute, Woods Hole, Massachusetts 02543, USA

THE presence of dissolved water can significantly change the sequence in which different minerals crystallize from a silicate magma^{1,3}. This will alter the compositional path followed by residual liquids (the 'liquid line of descent'), and modify the types and compositions of crystals that accumulate at the site of cooling. The sharp decrease in the solubility of water in silicate melts with decreasing pressure makes direct measurements of pre-eruptive concentrations in magmas difficult because much of the dissolved water boils off before eruption. Here we report experimental results on the crystallization of basalts that contain substantial dissolved water (up to 6 wt%), and show how the results can be used to infer the amount of dissolved H₂O involved in the petrogenesis of lavas that preserve a record of their liquid lines of descent, or the accumulated minerals left behind after solidification. The presence of abundant magmatic water is a signature of subduction-zone volcanism⁴, and can be used to associate a suite of magmas or igneous cumulates with a convergent-margin, back-arc or inter-arc tectonic setting.

Compositional variations observed in spatially and temporally associated volcanic rocks have traditionally been interpreted as a record of the conditions in which the parent magma crystallized. Similarly, the mineral associations and the compositional variations in the accumulated crystals preserve a record of the conditions of solidification. (Other magmatic processes that can lead to compositional variations in arc magmas include magma mixing and assimilation of crust and mantle. The influence of these other processes can be recognized and must be accounted for along with the effects of crystal fractionation.) The compositional variations in suites of island-arc basalts are often consistent with the early crystallization of olivine + clinopyroxene⁵, whereas the systematics of mid-ocean-ridge basalt suites point to olivine + plagioclase as the early crystallizing phases⁶. Olivine and high-Ca clinopyroxene are found together in igneous ultramafic cumulate rocks (wehrlites and olivine clinopyroxenites), and this association has been interpreted as an indication of crystallization at elevated pressure in the upper mantle^{7,8}. We use new experimental results (see ref. 9 for details) to show that

the signature of early olivine + high-Ca pyroxene crystallization in island-arc basalts results from magmatic H₂O contents of ~2–4 wt% at crustal pressures (6–10 km depth). We also show that the cumulates left behind by hydrous, low-pressure crystallization processes are distinct from those formed under anhydrous conditions at elevated pressures.

Melting experiments have established that primitive, sub-alkaline lavas, including both mid-ocean-ridge and high-alumina basalts, crystallize olivine ± plagioclase at crustal pressures under anhydrous conditions^{10,12}; clinopyroxene crystallizes in relatively evolved lavas. This crystallization sequence will accordingly produce olivine (dunite), olivine + plagioclase (troctolite) and olivine + plagioclase + clinopyroxene (olivine gabbro) cumulates. A few experimental studies have found olivine + clinopyroxene near the anhydrous liquidus at high pressures (~8–10 kbar)^{8,13,14}, leading to the common assumption that olivine + clinopyroxene cumulates form by crystallization in upper-mantle conditions.

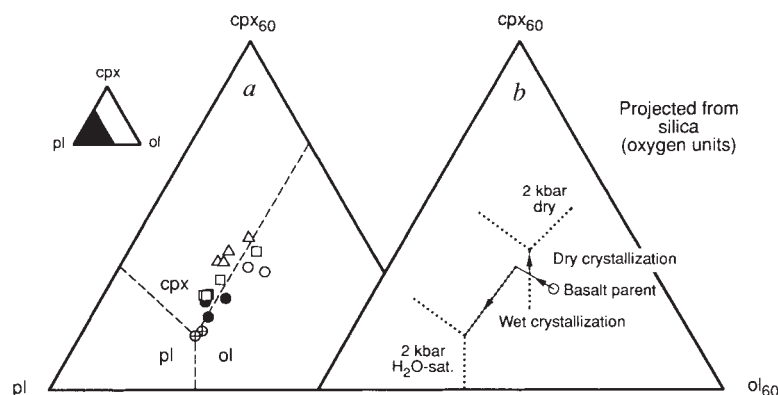
Water-saturated melting experiments performed on an (olivine + high-Ca clinopyroxene)-bearing basalt from ODP Leg 135 Hole 839B, in the Lau basin, produced an extended interval of olivine + clinopyroxene crystallization at a pressure of only 2 kbar (~6 km depth). This basalt was chosen for experimental study because it is fine-grained, equigranular and lacks complex phenocrysts. It therefore represents a liquid composition. This basalt crystallizes olivine + plagioclase near its liquidus at 1 atm,

TABLE 1 Electron microprobe analyses of experimental run products and xenolith clinopyroxenes

	1	2	3	4	5	6
SiO ₂	53.3 (2)	53.5 (1)	54.4 (1)	52.4 (4)	51.94	51.1
TiO ₂	0.60 (3)	0.73 (3)	0.70 (5)	0.33 (9)	0.36	0.48
Al ₂ O ₃	15.16 (8)	14.26 (7)	17.7 (1)	2.9 (3)	2.55	7.5
Cr ₂ O ₃	0.05 (3)	0.07 (3)	0.03 (2)	0.6 (1)	0.10	1.4
FeO*	8.85 (8)	9.4 (1)	8.0 (1)	6.4 (9)	4.12	4.9
MnO	0.22 (2)	0.19 (3)	0.22 (4)	0.18 (3)	0.21	—
MgO	9.36 (5)	7.91 (9)	5.85 (6)	16.9 (5)	16.22	17.9
CaO	11.3 (1)	11.5 (2)	10.8 (1)	20.5 (7)	24.82	14.4
Na ₂ O	1.48 (5)	1.55 (6)	1.85 (9)	0.14 (2)	0.20	1.0
K ₂ O	0.30 (2)	0.30 (1)	0.34 (1)	—	—	0.11
P ₂ O ₅	0.11 (3)	0.16 (2)	0.20 (3)	—	—	—
Total	100.73	99.57	100.09	100.35	100.52	98.79

Analyses given in wt%. 1. Sample 135–839B-15R-2, 63–67 cm melted at 1,210 °C, 1 atm⁹; 2. glass from sample 135–839B-15R-2, 63–67 cm after 18% crystallization at 1 atm under anhydrous conditions⁹; 3. glass from sample 135–839B-15R-2, 63–67 cm after 17% crystallization at 2 kbar under H₂O-saturated conditions⁹; 4. clinopyroxene coexisting with glass no. 3 from 2 kbar, H₂O-saturated melting experiment⁹; 5. clinopyroxene from low-pressure wehrlite xenolith from Adak island, Aleutian arc³¹; 6. clinopyroxene from high-pressure wehrlite xenolith from Black Rock Summit, Nevada, USA³². (Units in parentheses represent one standard deviation of least units cited, based on replicate analyses. Therefore, 53.3(2) should be read 53.3 ± 0.2.)

FIG. 1 a, Compositions of glasses from melting experiments performed at 2 kbar, under H₂O-saturated conditions (○: olivine+chromian spinel; ●: olivine+clinopyroxene+chromian spinel; ⊕: olivine+clinopyroxene+plagioclase+chromian spinel) and groundmass separates of olivine+clinopyroxene-phyric lavas from two volcanoes in the Vanuatu arc^{17,18} (□: Mere-lava; △: W. Epi) recast into mineral components⁴¹, and projected onto the pl-cpx-ol pseudo-ternary. Saturation boundaries at 2 kbar, H₂O-saturated conditions (---) are inferred from our experiments. b, Schematic liquid lines of descent for a primitive basalt (○) at 2 kbar, H₂O-saturated (path labelled Wet crystallization), and 2 kbar, anhydrous (path labelled Dry crystallization) at ~6 km depth. Saturation boundaries at 2 kbar, anhydrous, were estimated using the expressions of Kinzler and Grove⁴². (Subscripts on cpx and ol indicate fractional value of ternary shown as enlargement.)



anhydrous, with clinopyroxene joining the assemblage after 30 wt% crystallization⁹. Because H₂O expands the clinopyroxene stability field, at the expense of plagioclase, the olivine+clinopyroxene+plagioclase saturation boundary shifts to plagioclase-rich compositions^{9,15,16}. Therefore, H₂O-rich basalts crystallize olivine and olivine+clinopyroxene as they evolve to olivine+clinopyroxene+plagioclase saturation, producing dunites, wehrlites and gabbros.

The plagioclase-clinopyroxene-olivine (pl-cpx-ol) pseudo-ternary phase diagram in Fig 1a compares the compositions of groundmass separates of olivine-, clinopyroxene-phyric lavas from two volcanoes in the Vanuatu arc^{17,18} with glass from the melting experiments (2 kbar, H₂O-saturated). The Vanuatu lavas show olivine+clinopyroxene crystallization trends that are similar to the results of the experiments. The multiple saturation boundaries shown in Fig. 1b can be used to examine the changes in the liquid line of descent (LLD) resulting from H₂O dissolved in the melt, and the corresponding cumulate assemblages. Following olivine crystallization, the shift of the saturation boundaries caused by H₂O will lead to crystallization along the olivine+high-Ca clinopyroxene boundary. The proportions of clinopyroxene and olivine crystallizing at 2 kbar, H₂O-saturated conditions (78:22) are nearly identical to estimates for the Vanuatu lavas, as well as other (olivine+clinopyroxene)-bearing arc basalts^{16,18,19}. The ultramafic cumulates from this LLD will be dunite and wehrlite.

Figure 2a shows a global compilation of (olivine+clinopyroxene)-bearing arc basalts^{8,20,21} and a field for Aleutian lavas²²⁻²⁸, projected from ol onto the pl-cpx-SiO₂ pseudo-ternary phase diagram. Figure 2b shows the compositional path followed by a basalt during crystallization at 2 kbar under both hydrous and anhydrous conditions, similar to the paths shown in Fig. 1b. If the (olivine+clinopyroxene)-bearing arc lavas had crystallized under low-pressure anhydrous conditions, plagioclase would have been an early crystallizing phase. They

would have evolved along the anhydrous olivine+clinopyroxene+plagioclase boundary to produce residual liquids that are relatively depleted in Al₂O₃ and enriched in MgO and FeO. The global compilation of arc basalts shows that crystallization of olivine+clinopyroxene over an extended interval (~35 wt%) is necessary to explain the spread in compositions along the pl-cpx side of the pseudo-ternary phase diagram. This requires hydrous conditions to suppress the early crystallization of plagioclase. The Aleutian field, which includes basalts and andesites, extends the range of compositional variation along the hydrous, 2-kbar olivine+clinopyroxene+plagioclase saturation boundary¹⁶. Crystallization along this boundary will drive residual liquids away from the pl-cpx side of the pseudo-ternary diagram towards SiO₂. Based on our experiments and the extent of plagioclase suppression observed in the global arc-basalt data set, the pre-eruptive H₂O contents of the primitive (clinopyroxene)-bearing arc lavas were 2-4 wt%.

Table 1 compares the compositions of the liquids produced by ~17 wt% crystallization of the same starting composition (no. 1) at 1 atm, anhydrous (no. 2) and 2 kbar, H₂O-saturated (no. 3). The hydrous LLD, in which plagioclase crystallization is suppressed, produces a liquid that is enriched in Al₂O₃ and Na₂O and depleted in MgO and FeO compared with the anhydrous LLD (see also ref. 16). Because both melts have undergone the same amount of crystallization, the concentrations of the incompatible elements (TiO₂, K₂O, P₂O₅) are nearly identical, despite a difference in MgO content of ~2 wt%. This indicates that in subduction-related settings, where the abundance of H₂O dissolved in magmas varies widely¹⁶, differences in incompatible-element concentrations at a constant MgO value may reflect the differences between a hydrous and an anhydrous LLD, and therefore the presence or absence of plagioclase as an early crystallizing phase.

To assess the effect of magmatic H₂O on pyroxene composition, high-Ca clinopyroxenes produced at 1 atm (anhydrous),

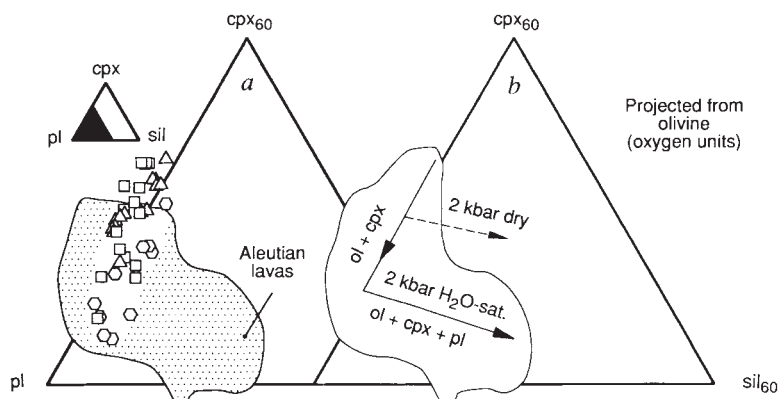
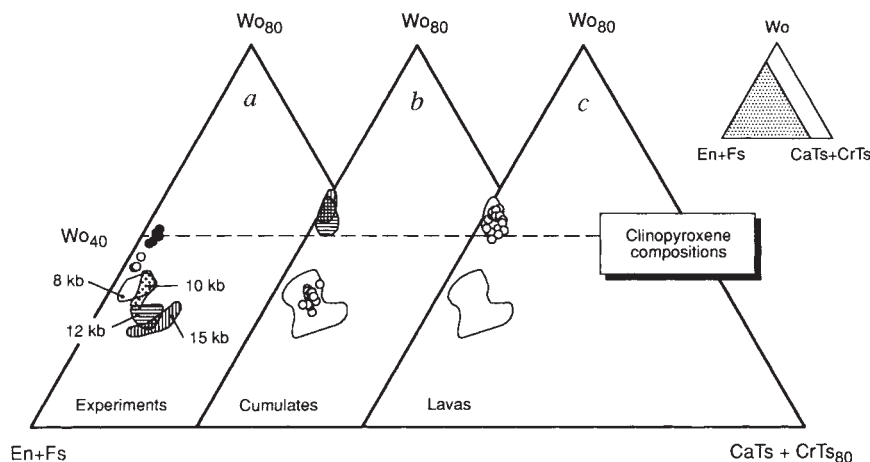


FIG. 2 a, Compilation of (olivine+clinopyroxene)-bearing lava compositions (△: Ambae in the Vanuatu arc²⁰; □: Grenada in the Lesser Antilles²¹; ○: Unalaska in the Aleutians⁸) projected from ol onto the pl-cpx-SiO₂ pseudo-ternary phase diagram. Also shown is the compositional field for basalts and andesites from the Aleutians²²⁻²⁸ compiled by Sisson and Grove¹⁶. b, Comparison of lava compositions shown in a with calculated olivine+clinopyroxene LLD and olivine+clinopyroxene+plagioclase saturation boundaries at 2 kbar, H₂O-saturated conditions (from our experiments⁹ and those of Sisson and Grove¹⁶), and at 2 kbar, anhydrous conditions (predicted using the method of Grove et al.¹²). Olivine+clinopyroxene LLD was calculated as described in ref. 9. (Subscripts on cpx and sil indicate fractional value of ternary shown as enlargement.)

FIG. 3 Clinopyroxene compositions recalculated into endmember components (En + Fs, (Mg, Fe)SiO₃; Wo, CaSiO₃; CaTs + CrTs, Ca(Al, Cr)SiAlO₆) and plotted on (En + Fs)–Wo–(CaTs + CrTs) ternary diagrams. *a* Clinopyroxenes produced in 1 atm, anhydrous (○), and 2 kbar, H₂O-saturated (●) melting experiments on olivine + clinopyroxene-phyric basalt from the Lau basin⁹. Also shown are fields representing clinopyroxenes produced in anhydrous melting experiments at 8 (open field), 10 (dotted field), 12 (field with horizontal lines) and 15 (field with vertical lines) kbar^{12,29,30} (experiments in which high-Ca pyroxene and low-Ca pyroxene coexist have been excluded). *b*, Clinopyroxenes from wehrlite and clinopyroxenite xenoliths from subduction (field with vertical lines)^{17,31} and non-subduction-related settings (○)^{32,33} and from olivine + clinopyroxene cumulates in ophiolites and ultramafic igneous complexes (field with horizontal lines)^{34–39}. *c*, Clinopyroxenes from (olivine + clinopyroxene)-bearing arc lavas^{17,18,20,39} (○). Open field represents experimental data shown in *a*. (Subscripts on Wo



and CaTs + CrTs indicate fractional value of ternary shown as enlargement.)

and 2 kbar (H₂O-saturated), are recalculated into pyroxene endmember components and plotted on the (En + Fs)–Wo–(CaTs + CrTs) ternary diagram in Fig. 3*a*. (The chemical composition of (En + Fs), Wo and (CaTs + CrTs) are given in Fig. 3 legend.) The effect of ~6 wt% H₂O dissolved in the melt is to increase the Wo content of the clinopyroxene, at nearly constant CaTs + CrTs. Compositional fields for low-TiO₂ (<0.9 wt%) clinopyroxenes produced in anhydrous melting experiments at 8–15 kbar^{12,29,30} are also shown. The effect of increasing pressure is to decrease the Wo content of clinopyroxene while increasing the CaTs + CrTs component. Therefore, cumulate wehrlites produced from an anhydrous melt at high pressures will contain clinopyroxenes with low Wo and high CaTs + CrTs, whereas those produced from a hydrous melt at crustal pressures will contain clinopyroxenes with high Wo and low CaTs + CrTs.

The ternary diagram in Fig. 3*b* compares clinopyroxenes from wehrlite and olivine clinopyroxenite xenoliths in sub-alkaline arc lavas^{17,31} with those from wehrlite and olivine-clinopyroxenite xenoliths in alkaline lavas erupted in non-subduction-related settings (US basin and range province³²; Hawaii³³). Also shown is a field for olivine + clinopyroxene cumulates from ophiolites, as well as Alaskan type and other ultramafic cumulate complexes^{34–39}. The clinopyroxenes from the arc xenoliths are characterized by high Wo contents and low CaTs + CrTs, indicating that they crystallized from a hydrous melt in the shallow crust. The Aleutian xenoliths contain amphibole and were recognized as the products of hydrous crystallization by Conrad and Kay³¹. The compositions of xenoliths from the non-subduction-related settings all occur within the high-pressure field defined by the experiments, suggesting anhydrous crystallization at 10–12 kbar (30–40 km depth). Table 1 compares the composition of clinopyroxenes produced experimentally under 2 kbar, H₂O-saturated conditions (no. 4), with xenolith clinopyroxenes whose compositions indicate crystallization at low pressure from a hydrous melt (no. 5), and at high pressure from an anhydrous melt (no. 6). The clinopyroxenes from the ophiolites and ultramafic intrusive complexes are coincident with the arc xenoliths, indicating that they also formed through hydrous crystallization at crustal pressures. This also suggests a subduction-related tectonic setting for these igneous bodies, consistent with other geochemical evidence^{31,39,40}.

The ternary diagram in Fig 3*c* shows the clinopyroxene phenocrysts from the (olivine + clinopyroxene)-bearing arc lavas^{17,18,20,39} shown in Figs 1 and 2. As expected, the clinopyroxenes are characterized by low CaTs and high Wo, and overlap the compositional fields for the hydrous experiments, arc xenol-

iths, ophiolites and ultramafic igneous complexes. The presence of these crystals in lavas that exhibit the hydrous LLD shown in Figs 1 and 2 is further evidence that the magmas parental to the wehrlite and olivine clinopyroxenite cumulates were compositionally similar to these magmas, and had similar abundances of dissolved magmatic H₂O. □

Received 24 February, accepted 29 July 1993.

- Holloway, J. R. & Burnham, C. W. *J. Petrology* **13**, 1–29 (1972).
- Kushiro, I. *Am. J. Sci.* **267A**, 269–294 (1969).
- Nicholls, I. A. & Ringwood, A. E. *J. Geol.* **81**, 285–300 (1973).
- White, D. & Waring, G. A. *Prof. Paper No. 440K*, 1–29 (US Geol. Survey, Washington DC, 1963).
- Perfit, M. R., Gust, D. A., Bence, A. E., Arculus, R. J. & Taylor, S. R. *Chem. Geol.* **30**, 227–256 (1980).
- Melson, W. G., Thompson, G. & van Andel, T. J. *H. J. geophys. Res.* **73**, 5925–5941 (1968).
- Elthon, D., Casey, J. F. & Komor, S. *J. geophys. Res.* **87**, 8717–8734 (1982).
- Gust, D. A. & Perfit, M. R. *Contr. Miner. Petrol.* **97**, 7–18 (1987).
- Gaetani, G. A., Grove, T. L. & Bryan, W. B. *Proc. ODP Sci. Res.* (in the press).
- Walker, D., Shibata, T. & DeLong, S. E. *Contr. Miner. Petrol.* **70**, 111–125 (1979).
- Grove, T. L., Gerlach, D. C. & Sando, T. W. *Contr. Miner. Petrol.* **80**, 160–182 (1982).
- Grove, T. L., Kinzler, R. J. & Bryan, W. B. *Mantle Flow and Melt Generation at Mid-Ocean Ridges* 281–310 (Am. geophys. Un., Washington DC, 1993).
- Bender, J. F., Hodges, F. N. & Bence, A. E. *Earth planet. Sci. Lett.* **41**, 277–302 (1978).
- Green, D. H., Hibberson, W. O. & Jaques, A. L. *The Earth: It's Origin, Structure, and Evolution* 265–299 (Academic, London, 1979).
- Yoder, H. S. *Carnegie Inst. Washington Yb.* **64**, 82–89 (1965).
- Sisson, T. W. & Grove, T. L. *Contr. Miner. Petrol.* **113**, 143–166, 167–184 (1993).
- Barsdell, M. J. *Petrology* **29**, 927–964 (1988).
- Barsdell, M. & Berry, R. F. *J. Petrology* **31**, 747–777 (1990).
- Foden, J. D. *J. Petrology* **24**, 98–130 (1983).
- Eggins, S. M. *Contr. Miner. Petrol.* **114**, 79–100 (1993).
- Thirwall, M. F. & Graham, A. M. *J. geol. Soc. Lond.* **141**, 427–445 (1984).
- Marsh, B. D. *J. Geol.* **84**, 27–45 (1976).
- Marsh, B. D. *Andesites* 99–114 (Wiley, New York, 1982).
- Kay, S. M. & Kay, R. W. *Contr. Miner. Petrol.* **90**, 276–290 (1985).
- Brophy, J. G. *Contr. Miner. Petrol.* **93**, 368–380 (1986).
- Myers, J. D., Marsh, B. D. & Sinha, A. K. *Contr. Miner. Petrol.* **94**, 1–11 (1986).
- Nye, C. J. & Reid, M. R. *J. geophys. Res.* **91**, 10271–10287 (1986).
- Debari, S., Kay, S. M. & Kay, R. W. *J. Geol.* **95**, 329–341 (1987).
- Bartels, K. S., Kinzler, R. J. & Grove, T. L. *Contr. Miner. Petrol.* **108**, 253–270 (1991).
- Grove, T. L., Kinzler, R. J. & Bryan, W. B. *Proc. ODP Sci. Res.* **106/109**, 9–17 (1990).
- Conrad, W. K. & Kay, R. W. *J. Petrology* **25**, 88–125 (1984).
- Wilshire, H. G. et al. *Prof. Paper No. 1443*, 111 (US Geol. Survey, Reston, Virginia, 1988).
- Wilkinson, J. F. G. *Contr. Miner. Petrol.* **58**, 181–201 (1976).
- Komor, S. C., Elthon, D. & Casey, J. F. *J. geophys. Res.* **90**, 7705–7736 (1985).
- Pallister, J. S. & Hopson, C. A. *J. geophys. Res.* **86**, 2593–2644 (1981).
- Ernewein, M., Pflumio, C., Whitechurch, H. *Tectonophysics* **151**, 247–274 (1988).
- Irvine, T. N. *Geol. Soc. Am. Memoir No. 138*, 240 (Geol. Soc. Am., Boulder, 1974).
- Snoke, A. W., Quick, J. E. & Bowman, H. R. *J. Petrology* **22**, 501–552 (1981).
- Sivell, W. & Rankin, P. *N.Z. J. Geol. Geophys.* **26**, 239–257 (1983).
- Elthon, D. *Nature* **354**, 140–143 (1991).
- Torney, D. R., Grove, T. L. & Bryan, W. B. *Contr. Miner. Petrol.* **96**, 121–139 (1987).
- Kinzler, R. J. & Grove, T. L. *J. Geophys. Res.* **97**, 6885–6906 (1992).

ACKNOWLEDGEMENTS. We thank F. Frey for comments on the manuscript, A. Saal and T. Wagner for discussions, M. Jercinovic for help with the electron microprobe, and T. Sisson for a review which helped to improve the manuscript. This work was supported by US National Science Foundation.

Rare species, the coincidence of diversity hotspots and conservation strategies

J. R. Prendergast^{*†}, R. M. Quinn^{*†}, J. H. Lawton^{*},
B. C. Eversham[†] & D. W. Gibbons[‡]

^{*} NERC Centre for Population Biology, Imperial College at Silwood Park, Ascot, Berkshire SL5 7PY, UK

[†] Biological Records Centre, NERC Institute of Terrestrial Ecology, Monks Wood, Abbots Ripton, Cambridgeshire PE17 2LS, UK

[‡] British Trust for Ornithology, The Nunnery, Nunnery Place, Thetford, Norfolk IP24 2PU, UK

SPECIES conservation *in situ* requires networks of protected areas selected for high conservation interest¹⁻³. Throughout most of the world, however, there are neither the resources nor the time to carry out detailed inventories for most taxa^{2,4} before designating protected areas. Site selection (on grounds other than availability) would be easier and more effective if two things were true: (1) habitats that are species-rich for one taxon are also species-rich for others⁵; and (2) rare¹ species occur in, and therefore benefit from the conservation of, species-rich habitats. Diversity (usually, species richness) and the presence of rare species are the most frequently cited criteria for site selection by conservationists⁶⁻⁸. Here, we use data on British plants and animals held by the Biological Records Centre (BRC)⁹ and the British Trust for Ornithology (BTO), mapped on a grid of 10 km × 10 km ('10 km squares') to examine the extent to which species-rich areas for different taxa coincide, and whether species-rich areas contain substantial numbers of rare species. The fine scale and high intensity of recording in Britain produces distributional datasets at least as good as and, in most cases, better than those available elsewhere. For Britain at least, we do not find strong support for either proposition. Species-rich areas ('hotspots'¹⁰) frequently do not coincide for different taxa, and many rare species do not occur in the most species-rich squares.

We used three groups of species: (1) birds (Aves), butterflies (Rhopalocera) and dragonflies (Odonata), 'popular', well studied taxa sometimes used as indicators¹¹, and which figure prominently in site selection; (2) liverworts (Hepaticae), inconspicuous 'lower' plants rarely considered in conservation planning; and (3) aquatic (angiosperm) plants, an ecologically (rather than taxonomically) defined group used as indicators of habitat quality¹². Similar analyses could be done with many other taxa, or taxonomic levels¹³; our analyses are illustrative, not exhaustive.

We define hotspots as the top 5% of record-containing 10-km squares (ranked by number of species per square, Table 1). 'Coldspots' are the most species-poor 5% of recorded squares. Many definitions of rarity exist^{14,15}; we use geographical distribution rather than abundance, because species with restricted distributions are most vulnerable to habitat loss or degradation¹⁶. We define rare species as those occupying <16 squares (the threshold for British Red Data Books¹⁷) and uncommon species as rare species plus those occupying 16-100 squares (also called 'scarce' species¹⁸). Data (Table 1, Fig 1a) were available from a maximum of 2,761 10-km squares, including the Isle of Man and Western Isles but excluding Shetland, Orkney and Ireland.

To what extent do hotspots for the five groups overlap? Most pairs of taxa exhibit nonrandom positive associations; maximum overlap (34%) is between butterflies and dragonflies (Table 2a). For two primarily southern taxa this association is weak, and it is even weaker for other pairwise comparisons. For example, only 12% of the bird and butterfly hotspots (two 'flagship' conservation taxa) coincide. Overlaps are especially low when groups have very different ecological requirements. Clearly, these data provide little support for the notion that areas species-rich for well known groups (birds or dragonflies) will also be species-rich for less well known taxa. There is also little overlap between coldspots (Table 2a) (at best 24% between the most depauperate areas for dragonflies and butterflies); indeed hot-spots for some taxa coincide with coldspots for others (Table 2b), albeit at a low level. Beyond pairwise comparisons, no single 10-km square is a hotspot for all five taxa. Only 2 squares are hotspots for 4 taxa and only 26 squares are hotspots for some combination of 3 taxa.

Although overlap of hotspots between taxa is low, a more optimistic picture emerges if overlap of hotspots with entire groups is considered. If it were possible to protect every hotspot for just one taxon, almost invariably more than half the species in the targeted taxon and of species in every other group would also occur in the protected areas (Table 2c); for example, 100% of butterfly species and over 90% of dragonflies, liverworts and aquatic plants occur in the set of 116 bird hotspots. Indeed the totality of bird hotspots contain more species of butterflies than do all the butterfly hotspots.

How well are restricted species represented in hotspots? Restricted species are more likely to be found in hotspots than in randomly selected squares or in coldspots (because hotspots contain more species in total), but surprisingly, many hotspots do not contain any rare species and about 17% do not contain any uncommon species (Table 3, Fig. 1b). Rare liverwort, aquatic plants and breeding birds frequently do not occur in any hotspots. One quarter or more of the uncommon species from

TABLE 1 Species occurrence data

	Butterflies	Dragonflies	Liverworts	Aquatic Plants	Breeding birds
Total number of species recorded	56	37	275	157	206
Number of species in most species-rich hotspot	44	28	137	88	109
Number of species in least species-rich hotspot	34	18	73	43	90
Number (proportion) of rare species	2 (0.04)	2 (0.05)	61 (0.22)	14 (0.09)	32 (0.16)
Number (proportion) of uncommon species	12 (0.21)	11 (0.30)	152 (0.55)	52 (0.33)	62 (0.30)
Number of hotspots	123	114	107	135	116
Number of coldspots	147	215	167	272	149
Total number of squares containing records	2,588	2,379	2,205	2,351	2,704

Data (presence in 10-km squares) for butterflies, dragonflies, liverworts and aquatic angiosperms were extracted from the BRC database. For these groups, only post-1960 records were used. Higher plants, butterflies and dragonflies have been extremely well recorded over the last 30 years, liverworts less so, but the distributions in Britain are now reasonably well documented. Alien and casual species are excluded from the dataset and subspecies treated as members of single aggregates where appropriate. The bird data are from the BTO Scottish Ornithologists' Club/Irish Wild Bird Conservancy 1988-91 Breeding Bird Atlas²² and include all confirmed breeding records and records of presence during the breeding season but without evidence of breeding and excluding introduced and feral species. Where the threshold for the 5% sets of hotspots and coldspots fell in a group with equal species-counts, it was raised to the next highest species counts, thus for butterflies the 5% set yields 129 hotspots with the threshold falling in a group of squares containing 33 species; the threshold was raised to 34 species, yielding 123 hotspots. We made no attempt to correct any of the data for variation in recording effort^{22,23} because distribution data used for site selection elsewhere in the world are also uncorrected. For definition of terms, see text.

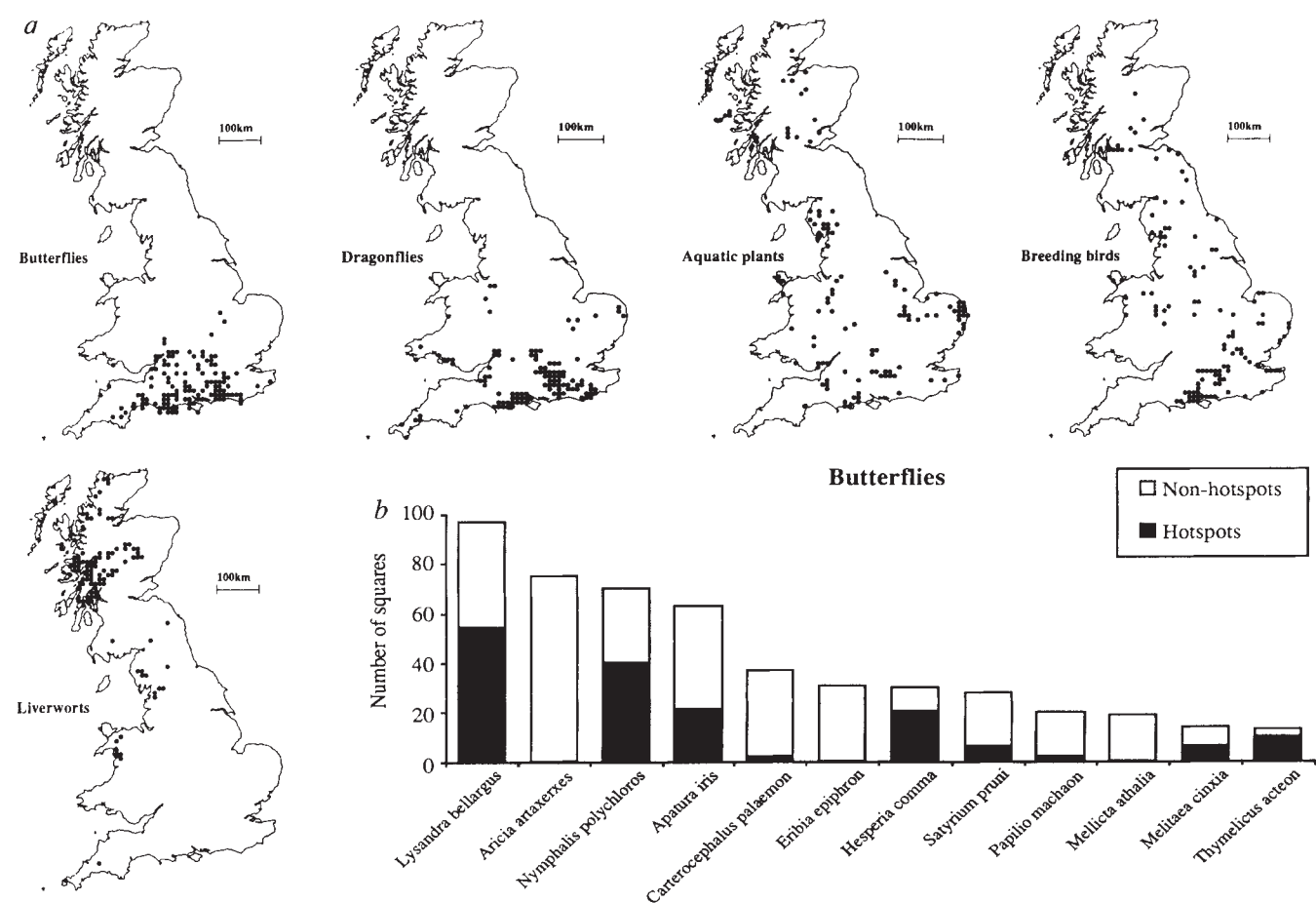


FIG. 1 a, Hotspots for representative UK taxa. Hotspots defined as the 10-km squares in the top 5 percentile of species richness within each group. Different taxa have markedly different distributions of hotspots, partly reflecting familiar features of their biology (for example butterflies

and dragonflies have concentrations of hotspots in the warmer, sunnier south of Britain, liverworts in the cooler, wetter north). b, The total number of squares and number of hotspots occupied by rare and uncommon species of butterflies.

four groups are not represented in hotspots. Sixteen per cent of rare, and 18% of uncommon bird species occur in coldspots but not in hotspots.

Several caveats are in order. Britain is an island with high geological and climatic heterogeneity, supporting habitats that have been modified and fragmented by man for millennia. British data, therefore, may not be representative of other areas or taxa but, in the absence of similar information from elsewhere, it would be unwise to ignore these provisional analyses. This study does not address the conservation of endemic species¹⁹, whether patterns of endemism in different taxa coincide, or the conservation of phylogenetically or genetically distinct taxa, or complementary biota^{3,20,21}. The British data are not perfect, although we doubt that better data would materially alter our conclusions.

Definitions of hotspots and coldspots and of rarity are arbitrary and our results are, of course, scale dependent. At coarser scales of resolution rare species and hotspots, and hotspots for different taxa, are more likely to appear coincident.

British data provide only weak support for the notion that selecting a limited number of conservation areas that are 'good' for one or two taxa, automatically includes species-rich areas for other taxa. A limited number of species-rich areas do not guarantee effective conservation for rare and restricted organisms, many of which occur outside species-rich areas. Elsewhere in the world, where distribution data are often sparse and mapped at a large scale, and where conservationists may be under pressure to make rapid decisions on reserve placement, we suggest that a strategy based solely on the two most popular

TABLE 3 Rare and uncommon species in hotspots and coldspots

	Hotspots					Coldspots				
	Butterflies	Dragonflies	Liverworts	Aquatic plants	Breeding birds	Butterflies	Dragonflies	Liverworts	Aquatic plants	Breeding birds
A	108 (0.88)	111 (0.97)	54 (0.50)	91 (0.67)	90 (0.78)	147 (1.0)	215 (1.0)	161 (0.96)	263 (0.96)	129 (0.87)
B	29 (0.24)	28 (0.25)	0 (0)	3 (0.02)	32 (0.27)	145 (0.99)	213 (0.99)	138 (0.83)	182 (0.67)	104 (0.70)
C	2 (1)	1 (0.50)	32 (0.52)	8 (0.57)	18 (0.57)	0 (0)	0 (0)	8 (0.13)	1 (0.07)	6 (0.19)
D	9 (0.75)	8 (0.73)	96 (0.63)	45 (0.87)	40 (0.65)	2 (0.17)	1 (0.09)	24 (0.16)	30 (0.57)	16 (0.26)

A, Number (proportion) of hotspots/coldspots containing no rare species; B, number (proportion) of hotspots/coldspots containing no uncommon species; C, number (proportion) of rare species occurring in hotspots/coldspots; D, number (proportion) of uncommon species occurring in hotspots/coldspots.

TABLE 2 Hotspot and coldspot overlap data

		Hotspots				
(a) Hotspot overlap and coldspot overlap						
	—	Butterflies	Dragonflies	Liverworts	Aquatic plants	Breeding Birds
Coldspots	Butterflies	—	0.34 (n = 99)	0.0 (n = 101)	0.6 (n = 118)	0.26 (n = 112)
	Dragonflies	0.24 (n = 89)	—	0.0 (n = 94)	0.25 (n = 111)	0.12 (n = 116)
	Liverworts	0.13 (n = 29)	0.09 (n = 145)	—	0.03 (n = 98)	0.23 (n = 116)
	Aquatic plants	0.13 (n = 111)	0.16 (n = 170)	0.11 (n = 142)	—	0.21 (n = 112)
	Breeding birds	0.15 (n = 115)	0.22 (n = 88)	0.14 (n = 94)	0.21 (n = 58)	—
(b) Hotspot overlap with coldspots						
Coldspots	Butterflies	—	0.00 (n = 99)	0.28 (n = 29)	0.01 (n = 111)	0.01 (n = 116)
	Dragonflies	0.02 (n = 108)	—	0.01 (n = 94)	0.02 (n = 111)	0.02 (n = 114)
	Liverworts	0.04 (n = 115)	0.07 (n = 107)	—	0.09 (n = 109)	0.01 (n = 104)
	Aquatic plants	0.08 (n = 118)	0.02 (n = 111)	0.11 (n = 98)	—	0.03 (n = 112)
	Breeding birds	0.0 (n = 115)	0.03 (n = 88)	0.07 (n = 94)	0.0 (n = 58)	—
(c) Taxon representation in hotspots						
	X↓ Y→	Butterflies	Dragonflies	Liverworts	Aquatic plants	Breeding birds
	Butterflies	0.91	0.91	0.64	0.82	1
	Dragonflies	0.82	0.92	0.72	1	0.92
	Liverworts	0.48	0.57	0.95	0.69	0.92
	Aquatic plants	0.86	0.88	0.91	0.96	0.94
	Breeding birds	0.7	0.73	0.75	0.88	0.87

a, Proportional overlap of hotspots with hotspots (top right) and hotspots with coldspots (bottom left) for each group, n = maximum possible number of overlaps, calculated as the smaller of the pair of hotspot groups adjusted for squares lacking records for either group. For example there are 123 butterfly hotspots and 107 liverwort hotspots; therefore the maximum possible overlap between these two groups is 107 squares. Of these, six liverwort hotspots lie in squares unrecorded for butterflies and are therefore excluded. Thus the actual maximum possible overlap for these groups is 101 squares. b, Proportional overlap of hotspots with coldspots for each group, n = maximum possible number of overlaps, calculated as above. Note that two values are shown for each pair of taxa. Thus, overlap between coldspots of butterflies with hotspots of liverworts is 0.28 but that of hotspots of butterflies with coldspots of liverworts is 0.04. c, Proportion of the species in the taxa listed in column X occurring in the full set of hotspots identified for the taxa in row Y (for example, 91% of the British butterfly species occur in the 123 squares identified as butterfly hotspots; the same 123 butterfly hotspots contain 82% of all the British dragonfly species, 48% of liverworts and so on).

criteria, diversity (species richness) and rarity, and on only one or a limited number of taxa, may fail to provide adequate protection for many other organisms. □

Received 30 April; accepted 12 July 1993.

- Soulé, M. (ed.) *Conservation Biology* (Sinauer, Sunderland, Mass., 1986).
- Groombridge, B. (ed.) *Global Diversity: Status of the Earth's Living Resources* (Chapman & Hall, London, 1992).
- Pressey, R. L., Humphries, C. J., Margules, C. R., Vane-Wright, R. I. & Williams, P. H. *Trends Ecol. Evol.* **8**, 124–128 (1993).
- Raven, P. H. & Wilson, E. O. *Science*, **258**, 1099–1100 (1992).
- Pearson, D. L. & Cassola, F. *Cons. Biol.* **6**, 376–391 (1992).
- Margules, C. & Usher, M. B. *Biol. Conserv.* **21**, 79–109 (1981).
- Usher, M. B. (ed.) *Wildlife Conservation Evaluation* (Chapman & Hall, London, 1986).
- Scott, M. J. et al. *Wildl. Monogr.* **123**, 1–41 (1993).
- Harding, P. T. & Sheail, J. *Biological Recording of Changes in British Wildlife* (ed. Harding, P. T.) 5–19 (HMSO, London, 1992).
- Myers, N. *The Environmentalist* **8**, 187–208 (1988).
- Kremen, C. *Ecol. Appl.* **2**, 203–217 (1992).
- Rorslett, B. *Aquat. Bot.* **39**, 173–179 (1991).
- Williams, P. H. & Gaston, K. J. *Biol. Conserv.* (in the press).
- Dury, W. H. *Biol. Conserv.* **6**, 162–169 (1974).
- Rabinowitz, D. *Biological Aspects of Rare Plant Conservation* (ed. Synges, H.) 205–217 (Wiley, New York, 1981).
- Thomas, C. D. & Mallorie, H. C. *Biol. Conserv.* **33**, 95–117 (1985).
- Perring, F. H. & Farrell, L. *British Red Data Books: 1, Vascular Plants* 2nd edn (RSCN, Nettleham, 1983).
- Stewart, A. & Pearman, D. *BSBI (Botanical Society of the British Isles) News* **57**, 17–21 (1991).
- International Council for Bird Preservation *Putting Biodiversity on the Map: Priority Areas for Global Conservation* (International Council for Bird Preservation, Cambridge, 1992).
- Vane-Wright, R. I., Humphries, C. J. & Williams, P. H. *Biol. Conserv.* **55**, 235–254 (1991).
- Crozier, R. H. *Biol. Conserv.* **61**, 11–15 (1992).
- Gibbons, D. W., Reid, J. B. & Chapman, R. A. *The New Atlas of Breeding Birds in Britain and Ireland: 1988–91* (in the press).
- Prendergast, J. R., Wood, S. N., Lawton, J. H. & Eversham, B. C. *Biodiversity Lett.* (in the press).

ACKNOWLEDGEMENTS. We thank the volunteer recorders who provided all the distribution data and S. Naem and R. I. Vane-Wright for helpful comments on the manuscript. Maps were produced by the DMAP program, written by Alan Morton, Imperial College. The Centre for Population Biology and the BRC are funded by the Natural Environment Research Council (NERC). The New Atlas of Breeding Birds in Britain and Ireland: 1988–91 (BTO) was funded initially by the Central Electricity Generating Board and subsequently by its four successors: National Grid Company, National Power, PowerGen and Nuclear Electric, additional funding was provided by the Nature Conservancy Council and subsequently by the Joint Nature Conservation Committee. J.R.P. is supported by a NERC studentship; R.M.Q. is supported by a NERC Special Topics Studentship from the Esmée Fairbairn Trust, both CASE awards with the Institute of Terrestrial Ecology.

Noise enhancement of information transfer in crayfish mechanoreceptors by stochastic resonance

John K. Douglass^{*†}, Lon Wilkens^{*}, Eleni Pantazelou[†] & Frank Moss^{*†§}

Departments of ^{*} Biology and [†] Physics, University of Missouri at St Louis, St Louis, Missouri 63121, USA

IN linear information theory, electrical engineering and neurobiology, random noise has traditionally been viewed as a detriment to information transmission. Stochastic resonance (SR) is a non-linear, statistical dynamics whereby information flow in a multi-state system is enhanced by the presence of optimized, random noise^{1–4}. A major consequence of SR for signal reception is that it makes possible substantial improvements in the detection of weak periodic signals. Although SR has recently been demonstrated in several artificial physical systems^{5,6}, it may also occur naturally, and an intriguing possibility is that biological systems have evolved the capability to exploit SR by optimizing endogenous sources of noise. Sensory systems are an obvious place to look for SR, as they excel at detecting weak signals in a noisy environment. Here we demonstrate SR using external noise applied to crayfish mechanoreceptor cells. Our results show that individual neurons can provide a physiological substrate for SR in sensory systems.

The concept of SR originated in efforts to explain the suggested periodicities in recurrences of the Earth's ice ages as the

§ To whom correspondence should be addressed.

† Present address: ARL Division of Neurobiology, University of Arizona, Tucson, Arizona 85721, USA.

result of stochastic and weak periodic forces acting in concert on a bistable, global climate model⁷⁻¹⁰. Subsequent experimental¹ and theoretical^{2,3} work demonstrated that the information content of a weak signal can be maximized by noise in certain nonlinear systems. An operational definition of SR is that some measure of the output coherence relative to the output noise passes through a maximum at an optimal value of the input noise.

As a preliminary search for SR in living systems, we have studied the timing of spiking events (action potentials) in single mechanoreceptor cells of the crayfish, *Procambarus clarkii*¹¹⁻¹³. As spiking can be induced by both noise and a coherent signal, we hypothesized that transmission of a weak mechanical stimulus could be enhanced by an optimal noise intensity. We were guided by an early review on statistical processes in neurons¹⁴ and by previous experiments involving external^{15,16} or internal¹⁷⁻¹⁹ noise which, in certain cases, was found to enhance some aspect of the quality of the response^{15,20}. In an experimental approach similar to the physical demonstrations of SR, we added an external noise source to a weak periodic signal, defining this combination as the stimulus. We obtained both power spectra (PS) and interspike interval histograms (ISIHS), which assess the coherence of the spiking activity with the signal frequency. An alternative measurement, the peri-stimulus time (cycle) histogram, which is sensitive to coherence with the signal phase is not considered here but is being studied in our laboratory.

Original studies of the ISIH of spontaneous and stimulated discharges in cat auditory fibres and their interpretation in terms of a noisy integrate-and-fire model demonstrated the exponential-like decay of peaks located at integer multiples of the stimulus period²¹. Although both single²²⁻²⁴ and populations²⁰ of integrate-and-fire models with noise have suggested²⁵ that neurons can exhibit a noise-enhanced response, these studies did not include a measure of the coherence relative to the noise, which would have been necessary to demonstrate SR. The nature of coherence in the ISIH in periodically forced, noisy, physical systems and its relation to physiological data from similarly forced sensory neurons¹⁹ has recently been discussed²⁶. Here we use two separate coherence measurements (the SNR_{PS} and the SNR_{ISIH} , defined below) to demonstrate SR experimentally in single neurons. These SNRs quantify the information content of the spike train without invoking any specific stimulus-encoding mechanism or neuronal model.

Our experiments used near-field mechanoreceptors, located on the crayfish tailfan, in which small motions of cuticular hairs are transduced by their associated sensory neurons into spikes that propagate centrally along the sensory nerves. In power spectra computed from the spikes (Fig. 1), the main feature due to the periodic signal is the narrow peak at the fundamental frequency riding on a broad noise background. This feature was enhanced by intermediate noise intensities (compare Fig. 1a, and b), but degraded by higher noise intensities (Fig. 1c). Interspike interval histograms (Fig. 2) were obtained at the same three external

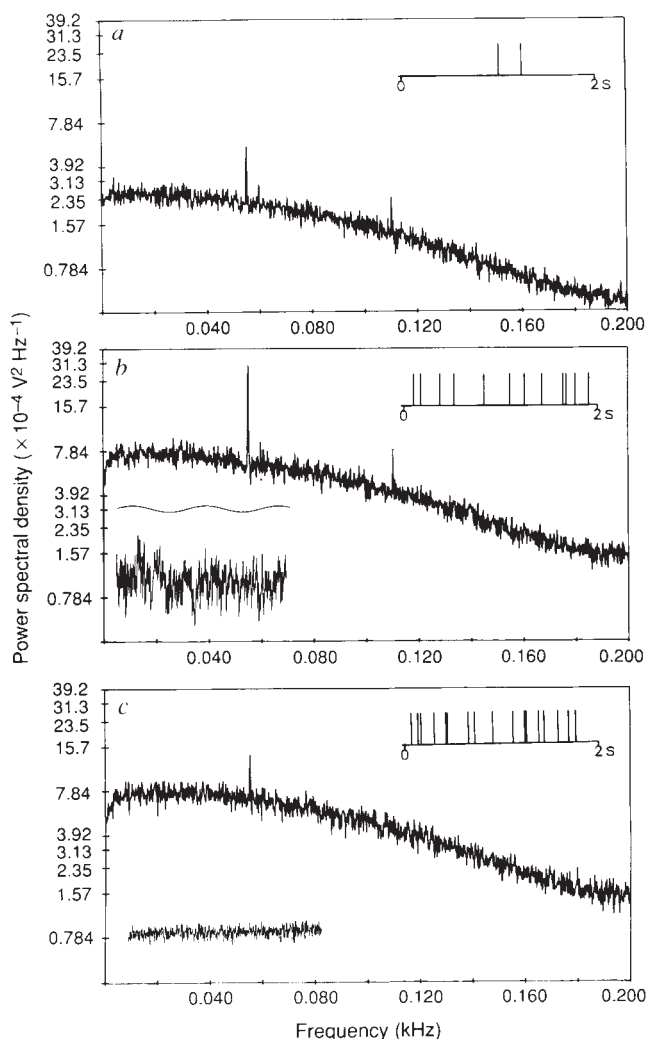


FIG. 1 Power spectra computed from the spiking activity of a crayfish mechanoreceptor, stimulated with a weak signal plus three external noise intensities: a, 0; b, 0.14; c, 0.44 V r.m.s. Insets at upper right show samples of the windowed spike trains from which the power spectra were computed. The total numbers of spikes used for the computations were 698 a, 2,390 b and 4,031 c. All spectra were acquired in the same amount of averaging time, 8.53 min. A single appendage (the telson) was isolated surgically from the crayfish tailfan, together with the associated nerve roots and terminal abdominal ganglion, and immersed in crayfish saline²⁹ at room temperature. The appendage was mounted vertically on an electromagnetic motion transducer activated by the sum of two inputs: a pure sinusoidal function of 55.2 Hz (the signal), plus a wide-bandwidth gaussian-distributed random function (the noise) as shown in the time series of the signal alone (upper left inset) (b) and the signal plus the noise (lower left inset) (b) and power spectrum (left inset) (c). The sample data shown in the insets were measured for the corresponding conditions stated above for a-c. The noise shown in the insets is quasi-white but the motion transducer introduced a roll-off of about 6.8 dB per frequency decade. The resulting stimulus was thus a combination of periodic and band-limited random motions of the hair relative to the saline, with the intensities of the noise and signal under independent control. Extracellular recordings from a nerve root were windowed to isolate spikes from a single sensory cell. The windowed events were shaped into 3.5-ms rectangular pulses, passed through a low pass (0–1.5 kHz) filter, and digitized at 10 kHz to obtain power spectra and ISIHS. To maximize the effects of the external noise, mechanoreceptors with low spontaneous spiking rates (low internal noise) were chosen for recording. The entire preparation was mounted on a vibration isolation table, the appendage was rotated so as to select the most effective stimulus direction, and a tuning curve was measured to select the most effective signal frequency. The signal intensity was then set slightly above that required for the minimum detectable SNR_{PS} , and various levels of noise were added to the stimulus. The noise and signal voltages delivered to the electromechanical transducer resulted in a stimulus velocity of $120 \mu\text{m s}^{-1}$ per V r.m.s. at 10 Hz as determined by an optical calibration. Crayfish mechanoreceptor cells are quite sensitive; without vibration isolation, the cells can easily detect weak building vibrations at 12.5 Hz.

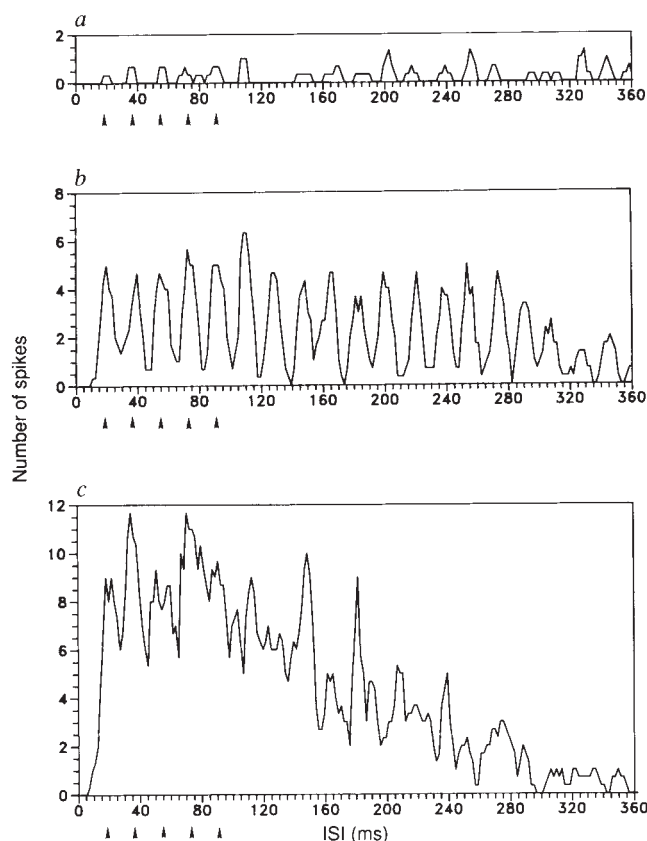


FIG. 2 Interspike interval histograms (ISI histograms) obtained under the same conditions used for Fig. 1. To emphasize the coherence of the ISI peaks with integer multiples of the stimulus period T_0 (18.1 ms), only intervals up to $20T_0$ are shown. Arrowheads mark the first five integer multiples of T_0 . For clarity these data were smoothed with a three-point moving average. The SNR_{ISI} calculations (Fig. 3) were obtained from unsmoothed data. The ISI histogram shown in *b* does not show the more familiar exponential-like decay of peak amplitudes beyond the first peak because the coherent portion of the stimulus is very weak. Under such conditions it is well known¹⁹ that the histogram is spread out over very long time intervals and that the first peak is not the one of maximum amplitude.

noise intensities. With no external noise (Fig. 2*a*) the spike rate was extremely low and most intervals were longer than those shown in Fig. 2. Intermediate noise (Fig. 2*b*) produced a striking growth of coherence marked by an increase in the amplitude of peaks located at integer multiples of T_0 , the stimulus period²⁶. Higher noise (Fig. 2*c*) resulted in increasing randomization of the peak structure.

Using various noise intensities, a series of power spectra such as those in Fig. 1 was used to calculate the SNR_{PS} by integrating the data in a small region around the fundamental peak to obtain the strength, S , the area under the signal peak above the noise. The SNR_{PS} , in decibels, was obtained using the standard definition: $\text{SNR}_{\text{PS}} = 10 \log_{10} [S/N(f_0)]$, where $N(f_0)$ is the amplitude of the broad-band noise background measured at the signal frequency f_0 . The results of these measurements (Fig. 3*a*) show a noise-induced signal enhancement of about 4.5 dB at an optimal noise intensity of ~ 0.14 V r.m.s.

As there is no formal definition of the SNR based on the ISI histogram, we developed an *ad hoc* definition assuming the occurrences of peaks at integer multiples of T_0 (ref. 26). The interspike intervals located around the first 100 peaks were summed on every interval $iT_0 \pm T_0/4$ for $1 \leq i \leq 100$, and called N_{max} . Similarly, the remain-

ing intervals located around the troughs were summed on $(i-1/2)T_0 \pm T_0/4$ and called N_{min} . We then defined the $\text{SNR}_{\text{ISI}} = 10 \log_{10} (N_{\text{max}}/N_{\text{min}})^2$. The results (Fig. 3*b*) parallel those computed from the power spectra: a noise-induced enhancement in the SNR_{ISI} but with a lower optimal noise intensity of ~ 0.10 V r.m.s.

Not all cells show SR, but we have measured at least some noise-induced enhancement in 10 out of the 11 cells tested. One cell, which had the largest internal noise, $N(f_0)$, showed only a monotonic decline instead of a maximum in the SNR.

Our experiment shows very clearly that weak signals can be enhanced by an optimal level of external noise in single sensory neurons. Does this mean that external noise can help crayfish detect weak signals which they could not detect in its absence? Experiments on sensory interneurons and/or behavioural studies will be necessary to answer this question. But a recent psychophysical experiment²⁷ and model²⁸ involving human perception of ambiguous figures in the presence of external noise suggest that this is so. □

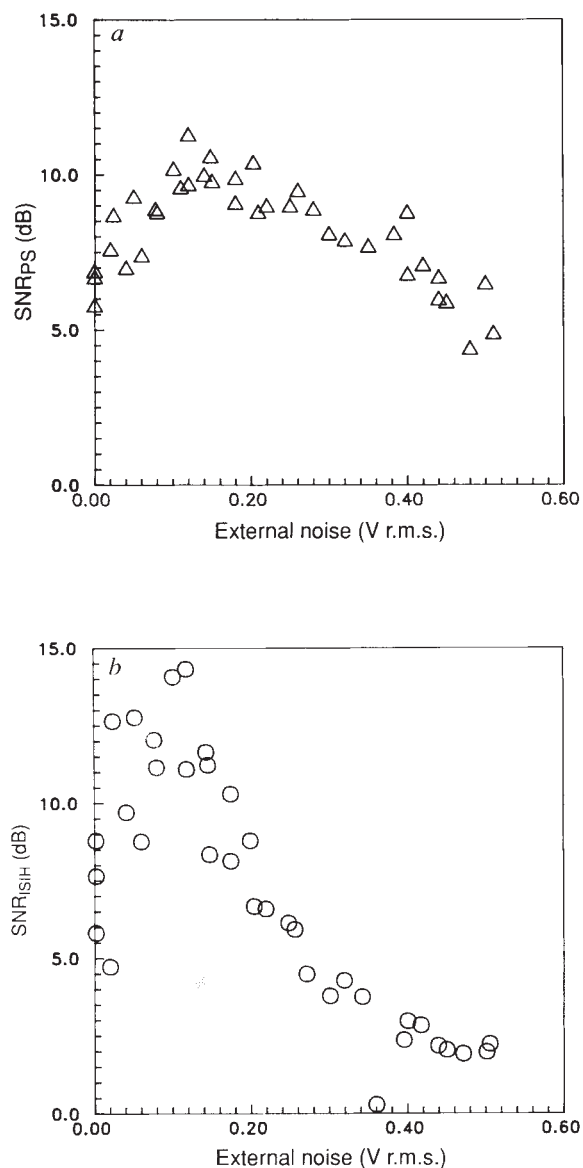


FIG. 3 Stochastic resonance measured from power spectra (*a*) and ISI histograms (*b*). The data shown in all figures were acquired from a single representative preparation.

Received 4 March; accepted 6 July 1993.

1. Fauve, S. & Heslot, F. *Phys. Lett.* **97A**, 5–7 (1983).
2. McNamara, B. & Wiesenfeld, K. *Phys. Rev. A* **39**, 4854–4869 (1989).
3. Jung, P. & Hänggi, P. *Europhys. Lett.* **8**, 505–510 (1989).
4. Gammaitoni, L., Marchesoni, F., Meneschi-Saetta, E. & Santucci, S. *Phys. Rev. Lett.* **62**, 349–352 (1989).
5. Moss, F., Bulsara, A. & Shlesinger, M. (eds) *J. stat. Phys.* **70**, 1–514 (1993).
6. Moss, F. in *An Introduction to Some Contemporary Problems in Statistical Physics* (ed. Weiss, G.) (SIAM, Philadelphia, in the press).
7. Benzi, R., Sutera, S. & Vulpiani, A. *J. Phys. A* **14**, L453–L457 (1981).
8. Nicolis, C. *Tellus* **34**, 1–9 (1982).
9. Benzi, R., Parisi, G., Sutera, A. & Vulpiani, A. *Tellus* **34**, 10–16 (1982).
10. Winograd, I. et al. *Science* **258**, 255–260 (1992).
11. Bush, B. M. H. & Laverack, M. S. in *The Biology of Crustacea* Vol. 3 (ed. Bliss, D. E.) 399–468 (Academic, New York, 1982).
12. Mellon, D. J. *exp. Biol.* **40**, 137–148 (1963).
13. Wiese, K. J. *Neurophysiol.* **39**, 816–833 (1976).
14. Moore, G., Perkel, D. & Segundo, J. A. *Rev. Physiol.* **28**, 493–522 (1966).
15. Bañó, W. Jr, Fuentes, J. & Segundo, J. *Biol. Cybern.* **31**, 99–110 (1978).
16. Narins, P. & Wagner, I. *J. acoust. Soc. Am.* **85**, 1255–1265 (1989).
17. Kaplan, E. & Barlow, R. B. Jr *Nature* **286**, 393–394 (1980).
18. Croner, L., Purpura, L. & Kaplan, E. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8128–8130 (1993).
19. Rose, J., Brugge, J., Anderson, D. & Hind, J. *J. Neurophysiol.* **30**, 769–793 (1967).
20. Knight, B. W. J. *gen. Physiol.* **59**, 734–766 (1972).
21. Gerstein, G. & Mandelbrot, B. *Biophys. J.* **4**, 41–68 (1964).
22. Stein, R. B. *Biophys. J.* **5**, 173–184 (1965).
23. Glass, L., Graves, C., Petrillo, G. & Mackey, M. C. *J. theor. Biol.* **86**, 455–475 (1980).
24. Glass, L. & Mackey, M. C. *J. math. Biol.* **7**, 339–352 (1979).
25. Knight, B. W. J. *gen. Physiol.* **59**, 767–778 (1972).
26. Longtin, A., Bulsara, A. & Moss, F. *Phys. Rev. Lett.* **67**, 656–659 (1991).
27. Chialvo, D. & Apkarian, V. *J. stat. Phys.* **70**, 375–392 (1993).
28. Ditzinger, T. & Haken, H. *Biol. Cybern.* **63**, 453–456 (1990).
29. van Harreveld, A. D. *Proc. Soc. exp. Biol. Med.* **34**, 428–432 (1936).

ACKNOWLEDGEMENTS. We thank A. Bulsara for discussions and C. Dames for help with experiments. This work was supported by the US Office of Naval Research.

Induction of human IgE synthesis in B cells by mast cells and basophils

Jean-François Gauchat*, Sybille Henchoz*, Gonzalo Mazzei*, Jean-Pierre Aubry*, Thomas Brunner†, Horst Blasey*, Paul Life*, Dominique Talabot*, Leopoldo Flores-Romo*, Jeff Thompson‡, Kenji Kishi§, Joseph Butterfield||, Clemens Dahinden† & Jean-Yves Bonnefoy*¶

* Glaxo Institute for Molecular Biology, 14 chemin des Aulx, 1228 Plan-les-Ouates, Geneva, Switzerland

† Inselspital Bern, 3010 Bern, Switzerland

‡ Glaxo Group Research, Greenford UB6 OHE, UK

§ Niigata University School of Medicine, Niigata 951, Japan

|| Mayo Clinic, Rochester, Minnesota 55905, USA

IMMUNOGLOBULIN E (IgE) is central to the induction of allergic diseases through its binding to the high-affinity receptor (FcεR1) on mast cells and basophils. Crosslinking by allergens of the bound IgE leads to the release of various inflammatory mediators. IgE production by B cells requires a physical interaction with T cells¹, involving a number of surface adhesion molecules^{1,2}, as well as the soluble factors interleukin-4 (IL-4)^{3,4} and IL-13 (ref. 5) produced by T cells^{6,7}, basophils⁸ and mast cells⁹. Here we report that, in the presence of IL-4, mast and basophilic cell lines can provide the cell contact signals that are required for IgE synthesis. The human cell lines HMC-1 (mast) and KU812 (basophilic) both express the ligand for CD40 (CD40L) which is shown to be responsible for the IgE production. Moreover, freshly isolated purified human lung mast cells and blood basophils are also shown to express CD40L and to induce IgE production. This evidence suggests that mast cells and basophils may therefore play a key role in allergy not only by producing inflammatory mediators, but also by directly regulating IgE production independently of T cells.

HMC-1 is an immature mast cell line expressing several mast cell antigens but not FcεR1 (ref. 10). KU812 is an immature basophilic cell line¹¹. To identify new cellular mechanisms for the control of IgE production, we tested the ability of these cell lines to induce IgE synthesis in purified B cells stimulated with IL-4 in the absence of T cells. Induction of IgE production and an increase in IgG synthesis was observed when B cells were incubated with HMC-1 or KU812 cells (Fig. 1). IgE production was preceded by a marked increase in IL-4-induced germ-line ε-RNA synthesis (measured at day 5, data not shown).

Studies using anti-CD40 monoclonal antibodies^{12,13} and

CD40L transfectants¹⁴ indicated that the CD40/CD40L interaction is a key event in the immunoglobulin switching that is required for IgE synthesis. Factors controlling IgE production regulate CD40L expression¹⁵ and T-cell-mediated stimulation of IgE synthesis can be inhibited by soluble forms of CD40 (ref. 16). Furthermore, direct evidence for the importance of CD40L in immunoglobulin switching was provided by the study of patients with hyper-IgM syndrome, which is linked to mutations in CD40L^{17–21}.

The induction of IgE and the increase in IgG synthesis brought about by HMC-1 or KU812 cells in the presence of IL-4 was inhibited by competition with a chimaeric recombinant fusion protein (Fig. 1) consisting of the extracellular domain of human CD40 fused to a murine IgG constant region (CD40-Fc). No inhibition was observed with the isotype-matched control antibody (Fig. 1). These results show that mast or basophilic cell lines could replace T cells to provide the contact-mediated signal that is required, in conjunction with IL-4, for the induction of IgE production. Inhibition of this production by CD40-Fc indicates an involvement of CD40L.

We investigated whether the expression of CD40L, thus far only detected on T cells²², could also be observed in HMC-1 or KU812 cells. We found that CD40L was indeed detectable by flow cytometry on both cell lines using fluorescein-labelled CD40-Fc as a probe (Fig. 2). Expression of CD40L was inducible by PMA and ionomycin in HMC-1 and constitutive in KU812 cells (Fig. 2). Expression in HMC-1 and KU812 cells of the messenger RNA coding for CD40L was studied by amplification (polymerase chain reaction) after reverse transcription (RT-PCR) and northern blot assays (Fig. 2). In unstimulated HMC-1, CD40L mRNA could only be detected by RT-PCR. The level of this RNA was maximal after stimulation of HMC-1 cells with calcium ionophore and phorbol ester for 3 hours (Fig. 2 and data not shown). In KU812 cells, there was a low level of constitutive expression of CD40L mRNA, a steady-state level sufficient to maintain detectable CD40L surface expression (Fig. 2). HMC-1 and KU812 CD40L mRNA were amplified after reverse transcription and cloned. Analysis of the sequence of the cloned, amplified complementary DNA, which included all the coding region, indicated that the mRNA coded for a protein identical to the CD40L identified in human T cells (data not shown). Therefore HMC-1 and KU812 cell lines express CD40L which, in conjunction with IL-4, is shown to be able to induce IgE production by B cells.

The immunosuppressive drug cyclosporin A (CsA) has recently been shown to be of therapeutic benefit in the treatment of certain atopic diseases^{23,24}. We therefore tested the effect of CsA on the up-regulation of CD40L mRNA in HMC-1 cells. When tested at doses similar to those used to demonstrate the effect of this drug on mast cell cytokine production²⁵, CsA abrogated the stimulation of CD40L mRNA expression induced by calcium ionophore and phorbol ester (Fig. 2c) which was paralleled by a decrease in the expression of the corresponding protein

¶ To whom correspondence should be addressed.

(data not shown). The inhibition of CD40L synthesis by mast cells might therefore contribute to the beneficial action of this drug.

To extend our observations found using cell lines to normal cells, histological detection of CD40L was done on skin tissue sections, lung mast cell suspensions, and blood basophils (Fig. 3). Using double staining, CD40L was shown to be expressed on c-kit ligand-positive mast cells from skin and lung as well as on blood basophils positive for FcεR1 α -chain. The physiologi-

cal significance of this CD40L expression was demonstrated in that highly purified lung mast cells and blood basophils were able to support IgE production by B cells, in the absence of T cells (Fig. 4). In contrast to mast cells, basophils were able to induce IgE synthesis even in the absence of exogenous IL-4. This agrees with the recent finding that basophils can secrete IL-4 (ref. 8). Mast cells which have also been reported to secrete IL-4 (refs 9, 26), may not have been adequately stimulated in the experimental system used here. The cell lines HMC-1 and

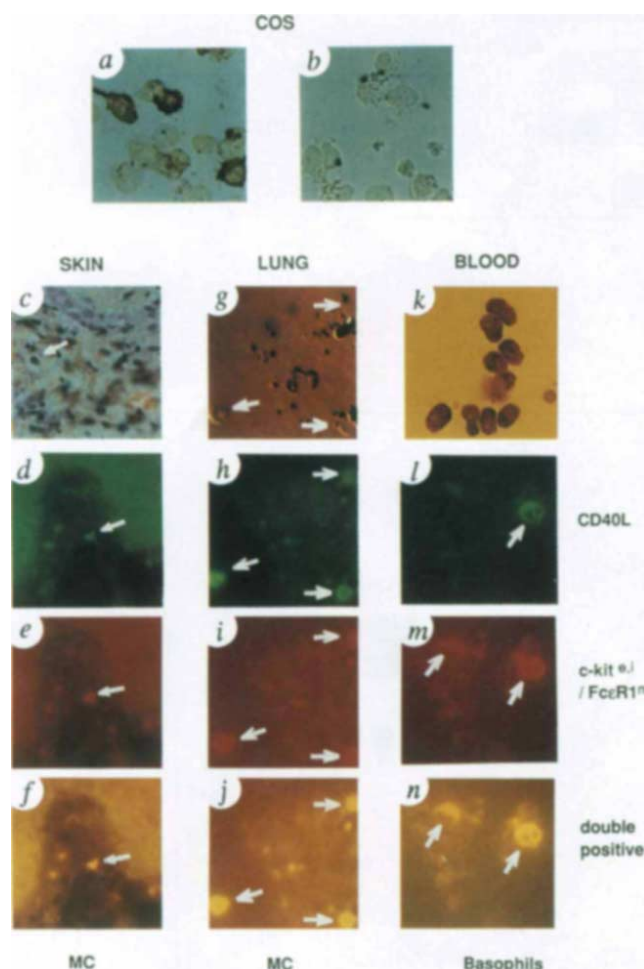


FIG. 3 Histological localization of CD40L on normal human mast cells and basophils. As a positive control for CD40L detection, CD40L-transfected (a) but not mock-transfected COS cells (b) were stained by CD40-Fc followed by peroxidase-labelled goat anti-mouse antibody and the mouse ClonoPAP-revelation system (both from Sternberger Monoclonals Incorporated, Baltimore, MA), or fluorescein-labelled goat anti-mouse antibody (Bioart, Meudon, France) (not shown) and analysed by microscopy (AxioPhot, Zeiss, Germany). Mast cells in the skin of a patient suffering from urticaria pigmentosa were stained in pink using toluidine blue (c). Granules of mast cells purified²⁹ from lung of a cancer patient are visualized by differential interference contrast (g). Purified blood basophils⁸ from healthy volunteers are stained by May-Grunwald-Giemsa (k). CD40L expression was studied by fluorescence microscopy on mast cells from tissue skin section (d and f), on mast cells from lung cell suspension (h and j) and on basophils from blood (l and n), using fluorescein-labelled CD40-Fc or IgG2a (control not shown). Mast cells were identified by biotinylated c-kit ligand (R&D systems, Minneapolis) revealed by phycoerythrin-labelled streptavidin (Becton-Dickinson) (e, f, i and j). Basophils were identified by biotinylated anti-FcεR1- α -chain monoclonal antibody (a gift of J.-P. Kinet) followed by phycoerythrin-labelled streptavidin (m and n). Mast cells (MC) are doubly stained for CD40-L and c-kit ligand (f and j). Some basophils are doubly stained for CD40L and FcεR1 α -chain (n). Overall magnification is $\times 260$ except for basophils (k-m) ($\times 410$).

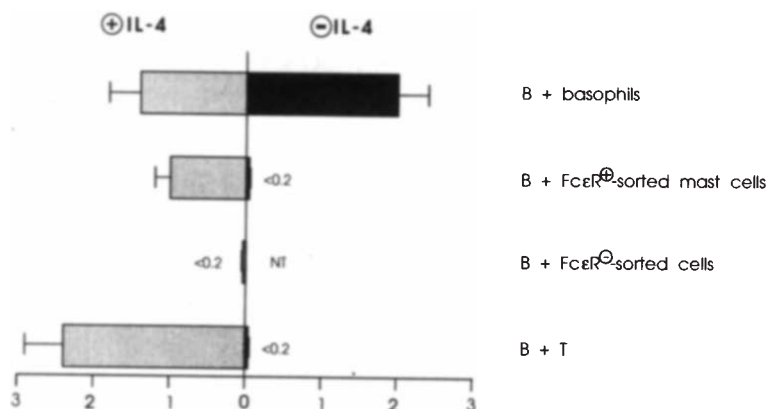


FIG. 4 Mast cells and basophils induce IgE by B cells. Secreted IgE was measured by enzyme-linked immunosorbent assay in supernatants of purified tonsil B cells (1×10^5 cells in 200 μ l) from sextuplet cultures incubated with IL-4 (200 U ml^{-1}) alone or IL-4 and the indicated cells for 12 days. A mast cell-enriched population was stimulated for 72 h with c-kit ligand²⁹ and FACS-sorted into FcεR1-positive and -negative populations using anti-FcεR1- α -chain monoclonal antibody followed by FITC-labelled goat anti-mouse antibody. Both populations (1×10^5) were tested but only the FcεR1-positive cells helped IgE synthesis and then only in the presence of exogenous IL-4. Blood basophils were enriched to 75% purity as described⁸ and highly purified (>97%) by negative selection by FACS using anti-CD3, CD20 and CD57 antibodies. Purified basophils were then stimulated for 3 h with PMA and ionomycin as described above. After washes, basophils (5×10^4) were incubated with B cells in the absence or presence of exogenous IL-4. Even in the absence of IL-4, basophils could induce IgE production. As a positive control, purified autologous T cells (1×10^5) were shown to induce IgE in presence of IL-4. NT Not tested. Data are presented for 1 out of 2 representative experiments.

KU812 were shown by northern blot analysis to express mRNA coding for cytokines which modulate IgE synthesis. HMC-1 expresses mRNA for IL-4, IL-6, tumour necrosis factor- α (TNF- α) but not IL-5 and interferon- γ (IFN- γ). KU812 expresses IL-5 and IL-6 but not IL-4, IFN α and TNF α mRNA (data not shown).

The data presented here demonstrate that IgE synthesis can be induced by the interaction of B cells with mast cells and basophils in the presence of IL-4. Mast cells and basophils can express CD40L which is known to be able to provide, in conjunction with IL-4, the minimal signals required to induce IgE production. This suggests that mast cells and basophils are likely to play an important role in the induction and maintenance of allergic reactions. Thus, immunoglobulin switching previously thought to take place only in lymph node germinal centres, may also occur in peripheral organs such as the lung or the skin. $\square$

Received 7 June; accepted 16 July 1993.

1. Vercelli, D., Jabara, H. H., Arai, K. I. & Geha, R. S. *J. exp. Med.* **169**, 1295–1307 (1989).
2. Aubry, J. P., Pochon, S., Graber, P., Jansen, K. & Bonnefoy, J. Y. *Nature* **358**, 505–507 (1992).
3. Coffman, R. L. et al. *J. Immunol.* **136**, 4538–4541 (1986).
4. Pene, J. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8166–8170 (1988).
5. Punnonen, J. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 3730–3734 (1993).
6. Howard M. et al. *J. exp. Med.* **155**, 914–923 (1982).

7. Minty, A. et al. *Nature* **362**, 248–250 (1993).
8. Brunner, T., Heusser, C. H. & Dahinden, C. J. *exp. Med.* **177**, 605–611 (1993).
9. Brown, M. A. et al. *Cell* **50**, 809–818 (1987).
10. Butterfield, J. H., Weiler, D., Dewald, G. & Gleich, G. D. *Leuk. Res.* **12**, 345–355 (1988).
11. Kishi, K. *Leuk. Res.* **9**, 381–390 (1985).
12. Jabara, H. H., Fu, S. M., Geha, R. H. & Vercelli, D. *J. exp. Med.* **172**, 1861–1864 (1990).
13. Gascan, H., Gauchat, J.-F., Aversa, G., Van Vlasselaer, P. & De Vries, J. E. *J. Immunol.* **147**, 8–13 (1991).
14. Armitage, R. J. et al. *Nature* **357**, 80–82 (1992).
15. Gauchat, J. F. et al. *FEBS Lett.* **315**, 259–266 (1993).
16. Fanslow, W. et al. *J. Immunol.* **149**, 655–660 (1992).
17. Aruffo, A. et al. *Cell* **72**, 291–300 (1993).
18. Korthäuer, U. et al. *Nature* **361**, 539–541 (1993).
19. DiSanto, J. P., Bonnefoy, J. Y., Gauchat, J. F., Fisher, A. & de Saint Basile, G. *Nature* **361**, 541–543 (1993).
20. Allen, R. C. et al. *Science* **259**, 990–993 (1993).
21. Fuleihan, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 2170–2173 (1993).
22. Spriggs, M. K. et al. *J. exp. Med.* **176**, 1543–1550 (1992).
23. Sowden, J. M. et al. *Lancet* **338**, 137–140 (1991).
24. Alexander, A. G., Barnes, N. C. & Kay, A. B. *Lancet* **339**, 324–328 (1992).
25. Burd, P. R. et al. *J. exp. Med.* **170**, 245–257 (1989).
26. Bradding, P. et al. *J. exp. Med.* **176**, 1381–1386 (1992).
27. Lederman, S. et al. *J. exp. Med.* **175**, 1091–1101 (1992).
28. Gauchat, J. F., Gauchat, D., de Weck, A. L. & Stadler, B. M. *Eur. J. Immunol.* **19**, 1079–1085 (1989).
29. Bischoff, S. C. & Dahinden, C. A. *J. exp. Med.* **175**, 237–244 (1992).

ACKNOWLEDGEMENTS. We thank I. Stamenkovic and J.-P. Kinet for the gifts of the CD40-IgG_{2a} construct and anti-Fc ϵ R1- α -chain antibody, respectively, T. Jomotte for technical assistance, G. Elson for recloning cytokine cDNA, M. Leemann-Husler for the graphics, J.-F. Nicoias for providing human skin biopsies, K. Ray for helpful discussions and K. Hardy, J. J. Mermoud and J. Knowles for their critical review of the manuscript. T. B. and C. D. were supported by the Swiss National Foundation.

Assembly and function of a quaternary signal transduction complex monitored by surface plasmon resonance

Stephan C. Schuster, Ronald V. Swanson,
Lisa A. Alex, Robert B. Bourret*
& Melvin I. Simon

Division of Biology, California Institute of Technology,
Pasadena, California 91125, USA

WE have used surface plasmon resonance biosensor technology to monitor the assembly and dynamics of a signal transduction complex which controls chemotaxis in *Escherichia coli*. A quaternary complex formed which consisted of the response regulator CheY, the histidine protein kinase CheA, a coupling protein CheW and a membrane-bound chemoreceptor Tar. Using various experimental conditions and mutant proteins, we have shown that the complex dissociates under conditions that favour phosphorylation of CheY. Direct physical analysis of interactions among proteins in this signal transduction pathway provides evidence for a previously unrecognized binding interaction between the kinase and its substrate. This interaction may be important for enhancing substrate specificity and preventing 'crosstalk' with other systems. The approach is generally applicable to furthering our understanding of how signalling complexes transduce intracellular messages.

Real-time analysis of specific protein-protein interaction is possible using surface plasmon resonance to monitor the changes in refractive index within a microflow cell^{1,2}. Internal to the flow cell is a dextran matrix covalently linked to a gold layer (Fig. 1). Carboxymethyl groups within the dextran layer provide sites for the covalent attachment of a ligand^{2–4} (Fig. 1). Binding of components to the immobilized ligand is measured as a change in refractive index and is directly proportional to the amount of bound protein⁴. This technique was used to study protein-protein interactions between components comprising the chemotaxis

signal transduction pathway in *E. coli*. In this system, CheA autophosphorylates⁵ on His 48 using ATP, and the phosphoryl group is then transferred to Asp 57 of the response regulator CheY^{6–8,23}. Phospho-CheY is believed to be the signal that results in clockwise rotation of the flagellar motor, inducing tumbling swimming behaviour in the bacterium⁹. Figure 2a shows the results of experiments involving the flow of different components of the phosphotransfer system over a surface containing immobilized CheY.

The rising slope of the resonance signal during the injection and the displacement of the baseline after the end of the injection indicate the binding of components to the surface. Wild-type CheA bound specifically to immobilized CheY51YC (Fig. 2a). In the control experiment no binding of CheA was observed with dextran matrices prepared in an identical manner, but without CheY in the immobilization step. Analysis of the binding and dissociation phases of this process revealed a fairly slow association rate of $368 \text{ M}^{-1} \text{ s}^{-1}$ and a dissociation rate of $1.14 \times 10^{-5} \text{ s}^{-1}$ (Fig. 3a, b). From these rate constants the K_d was calculated to be 30 nM, which suggests that these proteins would form stable complexes *in vivo*.

The kinase activity of CheA is modulated in a ligand-dependent manner through interaction with the CheW protein and membrane-bound chemoreceptors^{10–12}. Unliganded receptor activates the CheA kinase, stimulating the level of CheY phosphate formed while liganded receptor inhibits the kinase, resulting in low levels of CheY phosphate. When CheW and/or membrane vesicles containing the aspartate receptor (Tar) were tested for binding in the absence of CheA, they showed no affinity for the CheY51YC surface (Fig. 2a). However, when both Tar and CheW were preincubated with CheA, bound mass increased 4.5-fold relative to when the same amount of CheA alone was assayed (Fig. 2a). Stable ternary complexes between the receptor, CheA and CheW have been described¹³, suggesting that the increased binding is due to Tar/CheW/CheA complexes interacting with the CheY surface. Interestingly, if CheA was first bound to CheY, no further binding of CheW, Tar or a combination of the two components was detected (Fig. 2b). CheY binding to CheA alone may stabilize a form of the kinase that cannot productively interact with the receptor and CheW.

We probed the dynamics of the quaternary complex under conditions which would allow the phosphorylation of CheY. The complex was bound to CheY and the addition of 0.5 mM

* Present address: Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, North Carolina 27599-7290, USA.

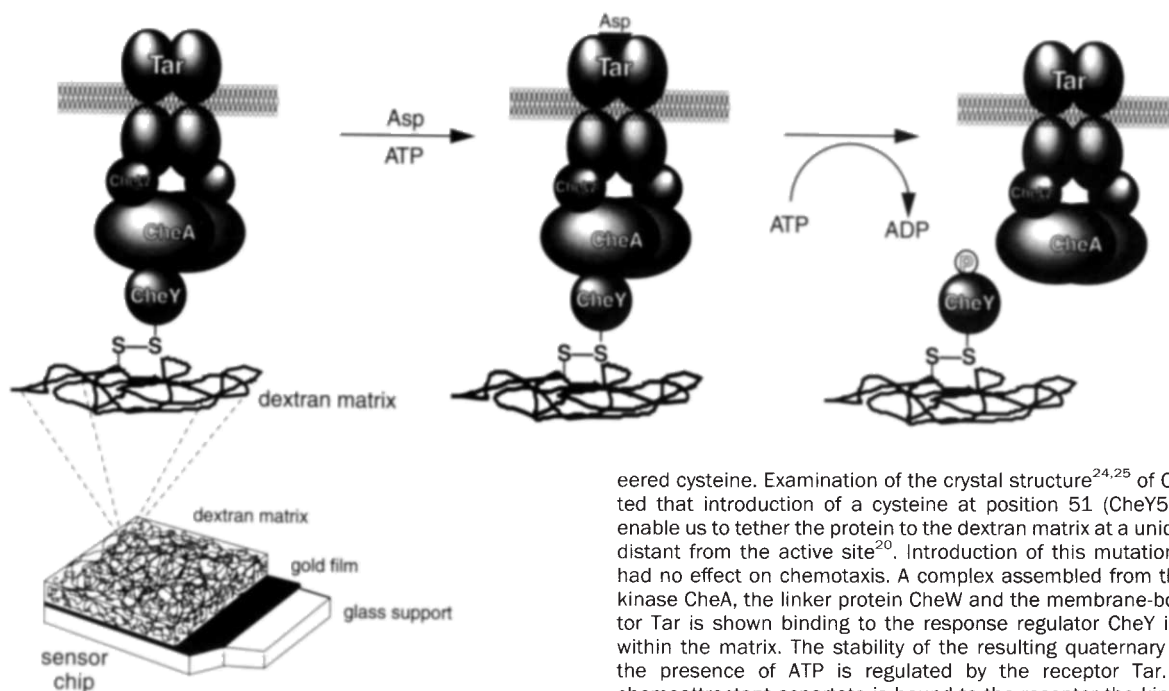


FIG. 1 Schematic drawing of the quaternary complex bound to the surface used to monitor the surface plasmon resonance signal. CheY is covalently attached to the dextran matrix of a BIAcore sensor chip consisting of a glass support covered by a gold layer to which the matrix is bound. CheY was immobilized through a disulphide bond to an engin-

eerred cysteine. Examination of the crystal structure^{24,25} of CheY indicated that introduction of a cysteine at position 51 (CheY51YC) would enable us to tether the protein to the dextran matrix at a unique position distant from the active site²⁰. Introduction of this mutation into CheY had no effect on chemotaxis. A complex assembled from the histidine kinase CheA, the linker protein CheW and the membrane-bound receptor Tar is shown binding to the response regulator CheY immobilized within the matrix. The stability of the resulting quaternary complex in the presence of ATP is regulated by the receptor Tar. When the chemoattractant aspartate is bound to the receptor the kinase is inactive in the presence of ATP, but when aspartate is not present the kinase is activated and can phosphorylate CheY, causing its release from the complex. The schematic representation of these reactions is supported by the experimental data in Fig. 2f. The response of the Tar/CheA/CheW complex to ligand is also controlled by methylation of specific glutamyl residues on the receptor¹⁴.

ATP resulted in a decrease of 30–70% (in a large series of experiments) of the mass bound to the surface (Fig. 2c). Thus, ATP promoted a >75-fold increase in the dissociation rate of CheY from the quaternary complex (Fig. 3c).

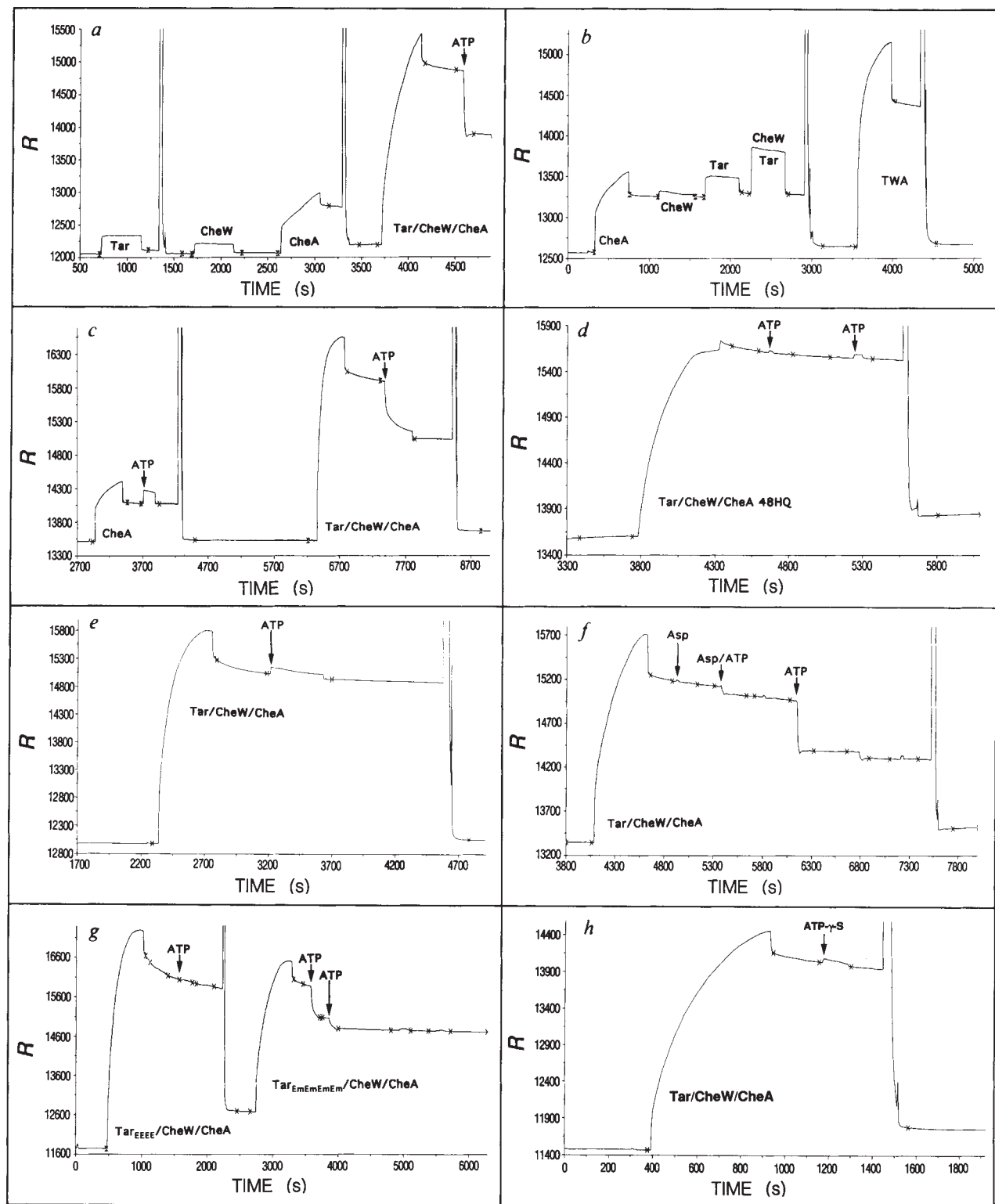
Gegner *et al.*¹³ demonstrated that the stability of the receptor/CheW/CheA complex in solution is not affected by the presence of CheY and ATP. This suggests that the decrease in mass observed with ATP treatment is due to the dissociation of the Tar/CheW/CheA ternary complex from CheY. A fraction of the mass remained on the CheY surface. This may be due to the presence of CheA that has bound CheY without receptor and CheW. Notably, although the Tar/CheW/CheA complex is stable, it is still in equilibrium with free CheA, which when bound to CheY is insensitive to ATP exposure (see below).

The release of Tar/CheW/CheA from CheY is specific for ATP, as other nucleotides had no effect on the stability of the quaternary complex. The non-hydrolysable ATP analogue ATP- γ S did not promote dissociation (Fig. 2h), which suggested that ATP binding alone does not result in the release but that hydrolysis is required. When quaternary complexes were formed using a mutant CheA (CheA48HQ) that cannot autophosphorylate⁵, they did not respond to ATP (Fig. 2d). In addition, quaternary complexes formed with a CheY mutant which lacks the phosphoryl acceptor site (CheY51YC57DA) could not be dissociated with ATP (Fig. 2e). This suggests that it is the phosphorylation of CheY which promotes release from the complex. Consistent with this is the fact that treatment of the CheY surface with acetyl phosphate, which is known to phosphorylate CheY at Asp 57 (ref. 22), diminishes its ability to bind CheA.

ATP-induced dissociation of the quaternary complex could be prevented if 1 mM aspartate is included with the ATP pulse (Fig. 2f). Aspartate alone had no effect on the stability of the complex (Fig. 2f). Closely related molecules such as asparagine

FIG. 2 Sensorgrams^{2–4} (plots of R versus time where R represents resonance units) of various chemotaxis proteins interacting with CheY51YC. **a**, The chemotaxis components Tar (provided as vesicles), CheW and CheA were passed over a CheY surface separately. Only CheA showed an observable interaction. When all three components were premixed in a test tube, a several-fold increase over the amount of binding that was seen by CheA alone was observed. **b**, CheA binding to CheY. CheW and/or Tar was introduced after CheA to assemble the quaternary complex with CheA. Neither component interacted with the preformed CheA/CheY complexes. When Tar, CheW and CheA were incubated before injection and then passed over the CheY surface, the increased response was observed. **c**, CheA/CheY complexes were treated with 0.5 mM ATP pulses lasting 2 min. No dissociation was observed. When the quaternary complex (Tar/CheW/CheA/CheY) was treated in an identical manner, 30–70% of the bound material dissociated, as seen by the decrease in resonance units. **d**, A quaternary complex assembled from Tar, CheW and the mutant protein CheA48HQ did not respond to ATP treatment. CheA48HQ was purified as described previously¹⁶. **e**, The ternary wild-type complex Tar/CheW/CheA bound to a surface prepared with the CheY mutant protein CheY51YC57DA is stable in the presence of ATP. The double mutant CheY protein was purified and coupled according to the procedure described for CheY51YC²⁰. **f**, The quaternary complex was assembled on the CheY surface to test the effect of liganded and unliganded receptor on the ATP-induced dissociation. Injection of 1 mM aspartate has no effect on the stability of the complex, as does 0.5 mM ATP in combination with 1 mM aspartate. Only when ATP alone was injected was a decrease in resonance units observed, suggesting that the ternary complex was released from CheY. **g**, Quaternary complexes assembled with fully demethylated receptor did not respond to ATP treatment, whereas complexes assembled with the fully methylated form exhibited release from CheY comparable to wild-type complexes after exposure to ATP. **h**, The quaternary complex is stable in the presence of ATP- γ S.

METHODS. Purified CheY51YC is immobilized within the flow-cell matrix



at an amount corresponding to 2,700R as described elsewhere²⁰. This corresponds to $\sim 2.7 \text{ ng mm}^{-2}$. All sensorgrams were recorded at a flow rate of $5 \mu\text{l min}^{-1}$ at 24°C in TEMK-buffer (50 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 5 mM MgCl_2 , 50 mM KCl), each injection of a protein component was $35 \mu\text{l}$, and all nucleotide injections were $5 \mu\text{l}$. Regeneration after each binding experiment was performed by injection of $2 \mu\text{l}$ 6 M guanidinium hydrochloride (injections that go off scale in the sensorgram, not indicated in each sensorgram separately) followed by extensive washing with TEMK-buffer. CheA and CheW were purified and membrane vesicles containing Tar were prepared as described¹⁴. Tar, referred to as wild-type, was prepared from a *cheR cheB* host. Reagents

were used at the following concentrations: $250 \mu\text{g ml}^{-1}$ CheA, $200 \mu\text{g ml}^{-1}$ CheW, $200 \mu\text{g ml}^{-1}$ Tar vesicles, 0.5 mM ATP, 10 mM acetyl phosphate and 1 mM L-aspartate. In cases where Tar/CheW/CheA complexes were preformed, the components were allowed to assemble for 1 min at room temperature before injection. In binding experiments where CheY51YC was omitted in the immobilization step, no binding of any of the other chemotaxis components was observed. Nucleotides that gave no effect on the release of the ternary complex from CheY included: GTP, ADP, GDP, ATP-γS, cAMP, dUTP and dTTP (all 0.5 mM).

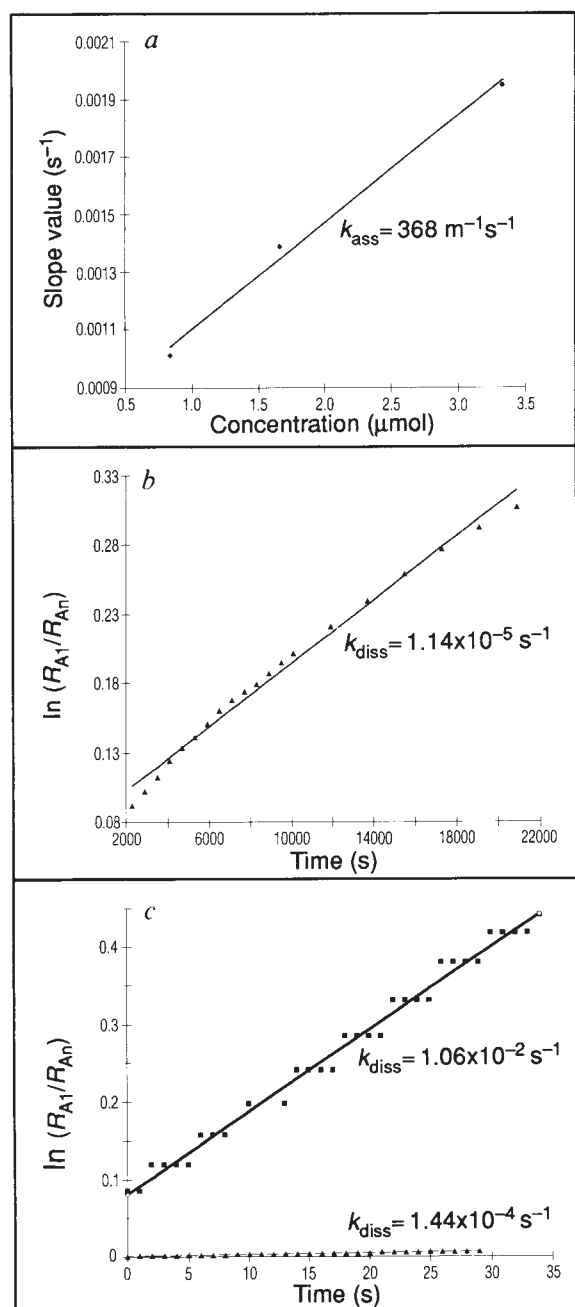


FIG. 3 Kinetic analysis of the interaction of CheA and Tar/CheW/CheA complexes with covalently immobilized CheY51YC. Purified chemotaxis proteins were used at various concentrations ranging from 0.83 μM to 6.67 μM in TEMK-buffer. They were allowed to interact with immobilized CheY51YC at a constant flow of 2 $\mu\text{l min}^{-1}$. The injection volume was 35 μl . The interaction curve was analysed in a R versus time plot. R values were continually calculated and displayed by the BIAlogue software (Pharmacia Biosensor). **a**, The slope of the interaction curve at each CheA concentration plotted against CheA concentration. The association constant k_{ass} was obtained from the slope of the fitted line²¹. **b**, Dissociation of CheA from CheY in continuous buffer flow. The dissociation rate (k_{diss}) was obtained from the slope of the $\ln(R_{A1}/R_{An})/(t_n - t_1)$ plot. The affinity constant of the CheA/CheY interaction was calculated²¹ from $K_d = k_{\text{diss}}/k_{\text{ass}}$. **c**, Comparison of the rates of dissociation of bound Tar/CheW/CheA from CheY in continuous buffer flow in the presence (■) and absence (▲) of ATP (0.5 mM). The presence of ATP accelerates the dissociation 75-fold. Because Tar was supplied as membrane vesicles, the exact concentration could not be determined, so the affinity of the Tar/CheW/CheA interaction with CheY is not given.

or succinate did not prevent ATP-induced dissociation. This result is consistent with biochemical studies on the regulation of phosphorylation which demonstrate activation of the kinase when the receptor is unliganded and inhibition of kinase activity when ligand is bound^{10,12,14}. Kinase activity is also modulated by the methylation state of the receptor. Covalent modification by reversible methylation of four glutamyl residues in the cytoplasmic signalling domain allows adaptation to new ligand concentrations¹⁵. Higher levels of methylation result in the increased ability of the receptor to stimulate kinase activity and hence produce large amounts of phospho-CheY^{12,14}. Quaternary complexes formed on the CheY surface containing Tar in a fully methylated state showed an ATP-inducible dissociation (Fig. 2g). When Tar in the fully demethylated form was used to assemble the complex it could not be dissociated with ATP (Fig. 2g). Thus, the behaviour of the differentially methylated receptors parallels the *in vitro* phosphorylation studies.

Our results show that CheY dissociation from the quaternary complex is correlated with conditions which favour phosphorylation. Previous studies have shown that CheA alone can phosphorylate CheY, albeit 300 times more slowly than the Tar/CheW/CheA complex^{5,10}. We did not observe any ATP-induced dissociation of CheA from CheY in the absence of Tar and CheW (Fig. 2c). This may be due to the slow rate of phosphorylation or to the non-equilibrium conditions within the flow cell. CheA is known to phosphorylate intramolecularly within a dimer^{16–18}. If the affinity between CheA monomers is weaker than the K_d of 30 nM that we measured for the CheA–CheY interaction, CheA may be bound to CheY as inactive monomers. Presumably, CheA initially binds to the CheY51YC surface as a dimer, but is then converted to a monomer by the continuous flow in the BIAcore apparatus. Tar, which is also a dimer¹⁹, may stabilize the dimeric form of CheA within the complex. Trapping of the monomeric form of CheA on CheY51YC could also account for the failure of the receptor and/or CheW to bind when CheA was first bound to CheY51YC (Fig. 2b), as the regulation of CheA autophosphorylation by receptor and CheW requires an intact CheA dimer¹⁶. An alternative possibility is that the release of phospho-CheY from CheA requires a conformational change induced by CheW and Tar.

The observation that CheA can bind stably to CheY and that this interaction can be monitored by surface plasmon resonance allowed analysis of the assembly, structure and dynamics of a quaternary signal-transducing complex. The use of receptors from crude membrane preparations for the assembly of various complexes including mutant proteins suggests that this approach should prove useful for other systems as well. The recognition of multicomponent signal-transducing complexes and their analysis by surface plasmon resonance should help us develop a deeper understanding of the nature of the specific protein–protein interactions that regulate the intracellular signalling circuitry. □

Received 23 April; accepted 1 July 1993.

1. Sambles, J. R., Bradbery, G. W. & Yang, F. *Contemp. Phys.* **32**, 173–183 (1991).
2. Jönsson, U. et al. *Biotechniques* **11**, 620–627 (1991).
3. Malmqvist, M. *Nature* **361**, 186–187 (1993).
4. Stenberg, E. et al. *J. Colloid Interface Sci.* **91**, 513–526 (1991).
5. Hess, J. F., Bourret, R. B. & Simon, M. I. *Nature* **336**, 139–143 (1988).
6. Hess, J. F., Oosawa, K., Kaplan, N. & Simon, M. I. *Cell* **53**, 79–87 (1988).
7. Wylie, D., Stock, A., Wong, C.-Y. & Stock, J. *Biochem. biophys. Res. Commun.* **151**, 891–896 (1988).
8. Hess, J. F. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7609–7613 (1987).
9. Bourret, R. B., Borkovich, K. A. & Simon, M. I. *Rev. Biochem.* **60**, 401–441 (1991).
10. Borkovich, K. A., Kaplan, N., Hess, J. F. & Simon, M. I. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1208–1212 (1989).
11. Borkovich, K. A. & Simon, M. I. *Cell* **63**, 1339–1348 (1990).
12. Ninfa, E. G., Stock, A., Mowbray, S. & Stock, J. *J. biol. Chem.* **266**, 9764–9770 (1991).
13. Gegner, J. A., Graham, D. R., Roth, A. F. & Dahlquist, F. W. *Cell* **70**, 975–982 (1992).
14. Borkovich, K. A., Alex, L. A. & Simon, M. I. *Proc. natn. Acad. Sci. U.S.A.* **89**, 6756–6760 (1992).
15. Springer, M. S., Goy, M. F. & Adler, J. *Nature* **280**, 279–284 (1979).
16. Swanson, R. V., Bourret, R. B. & Simon, M. I. *Molec. Microbiol.* **8**, 435–441 (1993).
17. Wolfe, A. J. & Stewart, R. C. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1518–1522 (1993).
18. Gegner, J. A. & Dahlquist, F. W. *Proc. natn. Acad. Sci. U.S.A.* **88**, 750–754 (1991).

19. Milligan, D. L. & Koshland, D. E. *J. biol. Chem.* **263**, 6268–6275 (1988).
20. Swanson, R. V., Schuster, S. C. & Simon, M. I. *Biochemistry* **32**, 7623–7629 (1993).
21. Karlsson, R., Michaelsson, A. & Mattsson, L. *J. Immun. Meth.* **145**, 229–240 (1991).
22. Lukat, G. S., McCleary, W. R., Stock, A. M. & Stock, J. B. *Proc. natn. Acad. Sci. U.S.A.* **89**, 718–722 (1992).
23. Sanders, D. A. et al. *J. biol. Chem.* **264**, 21770–21778 (1989).
24. Stock, A. M., Mottonen, J. M., Stock, J. B. & Schutt, C. E. *Nature* **337**, 745–749 (1989).
25. Volz, K. & Matsumura, P. *J. biol. Chem.* **266**, 15511–15519 (1991).

ACKNOWLEDGEMENTS. This work was supported by a grant from the US Public Health Service to M.I.S. S.C.S. is a recipient of a Max-Planck-Gesellschaft Award Fellowship, R.V.S. is supported by a National Research Service Award Fellowship and L.A.A. is supported by an American Cancer Society Fellowship.

SEC12 encodes a guanine-nucleotide-exchange factor essential for transport vesicle budding from the ER

Charles Barlowe & Randy Schekman*

Department of Molecular and Cell Biology, Howard Hughes Medical Institute, University of California, Berkeley, California 94720, USA

In yeast a type II integral membrane glycoprotein that is essential for transport vesicle budding from the endoplasmic reticulum (ER) is encoded by *SEC12* (refs 1–3). *SAR1* was discovered as a multi-copy suppressor of the *sec12-1^s* strain and encodes a GTPase of M_r 21,000 (21K) also essential for vesicle budding from the ER^{4,5}. *Sar1* is a peripherally associated membrane protein which shows enhanced membrane binding in cells containing elevated levels of Sec12 protein (refs 6, 7). We show here that a purified fragment of Sec12 promotes guanine-nucleotide dissociation from Sar1 whereas the purified mutant Sec12-1 has only 15% of the wild-type activity. GTP hydrolysis by Sar1 is not enhanced by Sec12, but is stimulated more than 50-fold by a mixture of Sec12 and Sec23, a GTPase-activating protein specific for Sar1 (ref. 8). We propose that Sec12 catalyses Sar1 guanine-nucleotide exchange in a process that recruits Sar1 to a vesicle formation site on the ER membrane.

Proteins transported between intracellular organelles and the cell surface are enclosed in transport vesicles that are targeted to and fuse with specific acceptor compartments. In transport from the ER to the Golgi apparatus, a set of interacting proteins has been identified that are required to produce functional transport vesicles from the ER membrane in a process requiring ATP and GTP^{9,10}. Small GTPases are thought to mediate vesicular transport in a cycle of GTP hydrolysis and nucleotide exchange that direct these processes in a unidirectional manner¹¹. Within the family of secretion-associated small GTPases, Sar1 is only 18–23% identical to the Ypt/Rab family and shares highest identity (35%) with yeast Arf1 (ref. 4).

Sec12 is a 70K glycoprotein with a prominent 40K N-terminal cytoplasmic domain, a single hydrophobic transmembrane domain and a glycosylated luminal domain⁷. Functional *SEC12* homologues have been identified in diverse species revealing conserved cytoplasmic domains whereas the luminal domains vary in size and amino-acid sequence¹². Expression of the cytoplasmic domain of Sec12 (Sec12 Δp) results in a soluble 40K protein that retards secretion in a wild-type strain and is toxic to a *sec12-1* strain. The toxic effect of Sec12 Δp can be compensated for *in vivo* by elevating the expression level of *SAR1* (ref. 7).

To examine the biochemical properties of Sec12 Δp we purified the protein to homogeneity from yeast after overexpression directed by the *GALI*-regulated promoter (see Fig. 2 legend for details). A reconstituted vesicle-budding reaction was used to

show that pure Sec12 Δp is a dose-dependent inhibitor of transport vesicle budding. Inhibition was maximal at a Sec12 Δp concentration of 1.6 μ M (Fig. 1a). Both the rate and extent of vesicle budding was inhibited by Sec12 Δp . Further, the inhibition was relieved by adding more purified Sar1 (Fig. 1), thereby reproducing the genetic interaction displayed by *SEC12* and *SAR1* *in vivo*.

We tested the capacity of the Sec12 fragment to modulate the GTP binding or hydrolysis properties of the Sar1 GTPase. Indeed, Sec12 Δp accelerated dissociation of guanine nucleotide from Sar1, as this activity co-purified with a single protein eluting from a CM (carboxymethyl)-Sephacrose column (Fig. 2). The peak fraction of Sec12 Δp was concentrated, and increasing amounts of the purified protein further stimulated the nucleotide dissociation rate of [³H]GDP bound to Sar1 (Fig. 3a). The calculated intrinsic off-rate for Sar1-GDP is 0.05 min⁻¹ whereas the presence of 0.3 μ M Sec12 Δp stimulated the rate to 0.40 min⁻¹. Sec12 Δp also stimulated the off-rate of [α -³²P]GTP bound to Sar1 from 0.10 min⁻¹ to 0.60 min⁻¹ in the presence of 0.3 μ M Sec12 Δp (not shown).

We tested the specificity of Sec12 Δp in stimulating nucleotide dissociation rates of other purified small GTPases preloaded with [³H]GDP. At a concentration of 0.5 μ M, Sec12 Δp did not affect the off-rates of [³H]GDP from Ypt1, ARF (ADP-ribosylation factor) or yeast Ras2 (Fig. 3b). In this experimental analysis, exchange factors for Ypt1 and ARF have not been isolated, although Cdc25, a known exchange factor for Ras2 (ref. 13), stimulated dissociation of GDP from Ras2 30-fold.

The *sec12-1* mutation has been identified as a C→T transition causing a P73L mutation within the cytoplasmic domain of

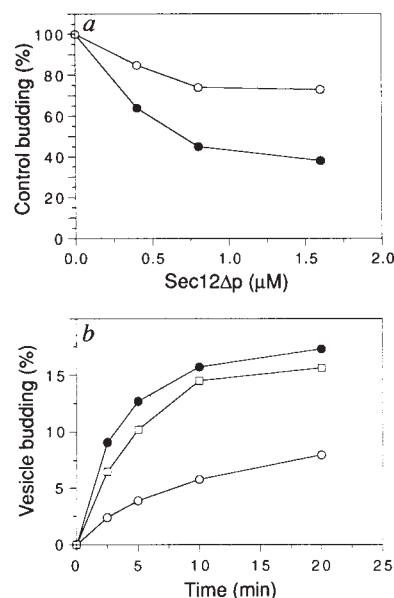


FIG. 1 Sec12 Δp inhibits the rate and extent of transport vesicle formation from the ER. *a*, Vesicle formation from ER membranes measured in the presence of wild-type cytosol and increasing amounts of Sec12 Δp (●) or wild-type cytosol, and increasing amounts of Sec12 Δp and 0.2 μ M of purified Sar1 (○). Control budding represents 16% packaging of total core-glycosylated pro- α -factor into transport vesicle intermediates after 20 min. *b*, A time course of vesicle budding was done with wild-type cytosol (●), in the presence of 0.6 μ M Sec12 Δp (○) or 0.6 μ M of Sec12 Δp and 0.2 μ M of Sar1 (□).

METHODS. Assays were done²² in the presence of 10 μ g of membranes and 200 μ g of cytosol prepared from strain RSY607 in a total reaction volume of 50 μ l at 29 °C. Sec12 Δp (see Fig. 2 legend) and Sar1 were purified as described⁵.

* To whom correspondence should be addressed.

Sec12 (ref. 7). This mutation was transferred to the *GALI* expression construct, allowing overproduction of Sec12 Δ p-P73L. The mutant protein was stably expressed at 25 °C and purified with chromatographic properties identical to Sec12 Δ p. Sec12 Δ p-P73L exhibited low activity in the exchange assay at 25 °C (not shown).

We used an independent assay for Sec12 Δ p interaction with Sar1 by coupling GTP hydrolysis with nucleotide exchange. Purified Sec23 has been shown specifically to activate Sar1

GTPase 15-fold⁸. Thus, we reasoned that in the presence of Sec23, dissociation of nucleotide from Sar1 would become limiting. As shown in Fig. 4a, a cycle of guanine-nucleotide hydrolysis and exchange demonstrates the reconstitution of this reaction with purified Sar1, Sec23 and Sec12 Δ p. Mixture of Sec12 Δ p and Sar1 results in an insubstantial increase in GTP hydrolysis, indicating that the hydrolytic step of this cycle is rate-limiting under our conditions. But addition of both Sec12 Δ p and Sec23 stimulated the rate of GTP hydrolysis more than 50-fold. Purified Sec12 Δ p-P73L also increased overall hydrolysis when added at an equimolar concentration to the wild-type protein, although the level of stimulation is only 15% of Sec12 Δ p (Fig. 4a, column 8). Sec12 Δ p-P73L did not display temperature sensitivity in the coupled GTPase assay (not shown). Affinity-purified polyclonal antibodies against Sar1 specifically inhibit Sar1 function in an *in vitro* vesicle budding assay⁵. These antibodies also inhibited the coupled hydrolysis reaction (Fig. 4a, column 9) whereas preimmune IgG had no effect.

Figure 4b shows a model for the cycle of GTP hydrolysis and exchange by Sar1, where transport vesicle budding is modulated by the integral membrane protein Sec12 and the peripherally associated Sec23 complex. Soluble nucleotide exchange factors have been identified for small GTPases¹³⁻¹⁵ and ATPases^{16,17}, but the transmembrane character of Sec12 is an interesting feature for this nucleotide exchange factor and is reminiscent of signal transduction pathways. Sec12 could itself be regulated by other proteins within the lumen, or on the cytosolic face of the ER to initiate vesicle budding at a specific membrane site. Although the activation of specific guanine-nucleotide exchange

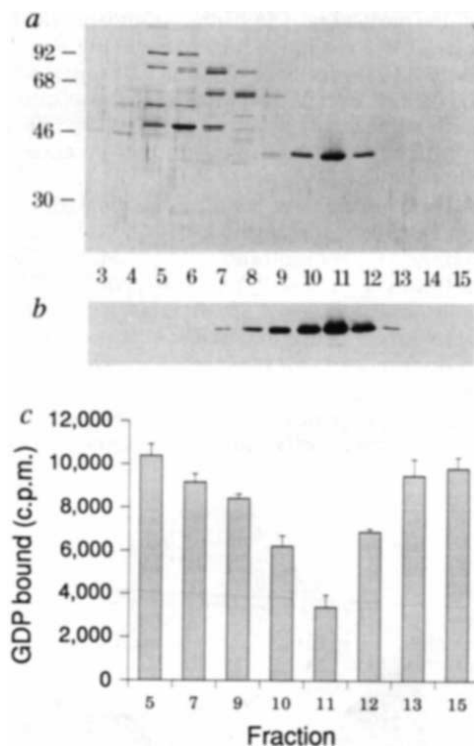


FIG. 2 Purified Sec12 Δ p stimulates dissociation of GDP from Sar1. Analysis of CM-Sepharose fractions from the final purification step of Sec12 Δ p: a, silver-stained SDS-PAGE gel; b, immunoblot with anti-Sec12 antibodies; and c, [³H]GDP bound to Sar1 after a 4-min incubation in the presence of competitor GTP and each column fraction, as determined by a filter binding assay⁵.

METHODS. Sec12 Δ p was purified from strain YPH500 harbouring plasmid pCEY4 described previously^{7,23}. Yeast (1 l) was grown in selective minimal medium and 2% raffinose to an A_{600} of 1.5. Cells were transferred to 2 l of rich medium (YP) containing 2% galactose and grown for 12 h to an A_{600} of 6.0. Cells were collected, washed with water and resuspended in 25 ml buffer A (25 mM K-HEPES, pH 6.8, containing 150 mM CH₃COOK, 0.25 M sorbitol and 1 mM magnesium acetate). The remaining steps were done at 4 °C. Cells were agitated in a 20-ml suspension of glass beads to achieve 50% cell lysis. A 100,000g supernatant fraction was prepared from this crude lysate and fractionated on a 600-ml 4.5 × 55 cm Sephacryl S-100 column developed with buffer A containing 10 mM CH₃COOK. The peak fractions containing immunoreactive Sec12 Δ p were pooled (50 ml) and loaded onto a 100-ml DEAE-Sepharose CL-4B column. The column was washed with 500 ml of buffer A containing 10 mM CH₃COOK and proteins were eluted with a 500-ml gradient from 10 to 400 mM CH₃COOK in buffer A. The peak of immunoreactive protein was pooled (75 ml), diluted with an equal volume of water and applied to a 10-ml CM-Sepharose column equilibrated with buffer A containing 100 mM CH₃COOK. This column was washed with 100 ml of equilibration buffer, and protein was eluted with a 200-ml gradient from 0.1 to 1.75 M CH₃COOK in buffer A. Fractions (5 ml) were collected as Sec12 Δ p eluted at a position corresponding to 1.0 M CH₃COOK. Column fractions (5- μ l aliquots) were analysed by SDS-PAGE²⁴ and immunoblot²⁵ or for an effect on Sar-[³H]GDP dissociation rate.

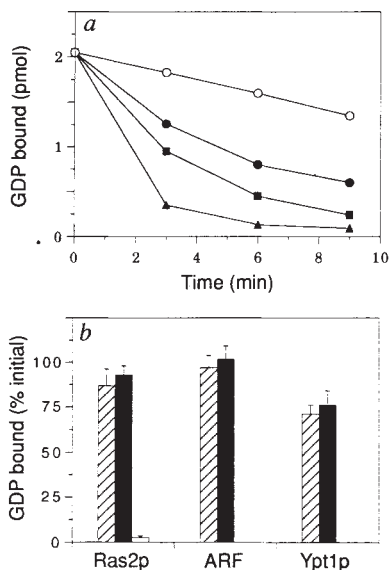


FIG. 3 Effects of Sec12 Δ p on the dissociation kinetics of prebound [³H]GDP from Sar1 and other small GTPases. a, Sar1-[³H]GDP (0.07 μ M) was incubated with 0.5 mM competitor GTP (○) and 0.03 μ M (●), 0.10 μ M (■) or 0.30 μ M (▲) Sec12 Δ p. The remaining amount of bound nucleotide was determined at various times by the filter binding assay. b, Percentage of initial [³H]GDP bound to Ypt1, ARF and Ras2 (~0.2 μ M each) was determined after 30 min in the absence (solid bars) or presence of 0.5 μ M Sec12 Δ p (hatched bars) and 0.5 mM competitor GTP. The addition of Cdc25 stimulated nucleotide dissociation of [³H]GDP from Ras2 (open bar).

METHODS. Sar1 was loaded with [³H]GDP⁵ and diluted into a 30- μ l reaction containing 0.5 mM GTP, 25 mM K-HEPES (pH 6.8), 0.1% Triton X-100 and 1 mM magnesium acetate. Exchange assays were done at 25 °C and nucleotide bound to a GTPase was quantified after retention on a nitrocellulose filter⁵. The purification and pre-binding of nucleotide to recombinant Ypt1 (ref. 26), ARF²⁷ and Ras2 (ref. 28) have been described elsewhere.

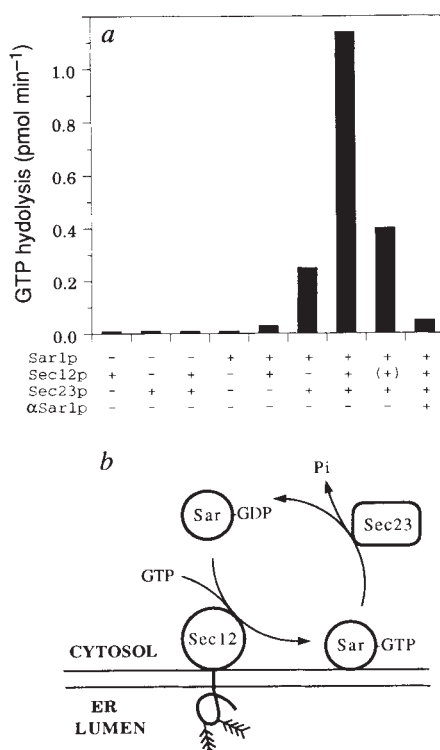


FIG. 4 Sec12 Δ p and Sec23 together stimulate Sar1 GTPase activity. **a**, Various combinations of Sec12 Δ p (0.5 μ M), Sec23 (0.5 μ M) and Sar1 (0.5 μ M) were incubated in 25 mM K-HEPES (pH 6.8), 0.1% Triton X-100, 1 mM magnesium acetate and 5 μ M [α -³²P]GTP at 25 °C for 10 min. GDP was quantified after separation of reaction products on polyethyleneimine cellulose plates. Sec12 Δ p-P73L was also added at 0.5 μ M, designated (+) in column 8. Affinity-purified anti-Sar1 antibodies⁵ were added to a final concentration of 0.1 mg ml⁻¹. **b**, Model for the cycle of GTP binding and hydrolysis by Sar1 regulated by Sec23 and Sec12.

factors by G-protein-coupled receptors or receptor tyrosine kinases presents an appealing model for Sar1 activation, there is no evidence for a ligand-induced stimulation of vesicle traffic from the ER.

The small GTPase ARF, involved in intra-Golgi transport, is also recruited to Golgi membranes in a process involving nucleotide exchange^{18,19}. The ARF-specific exchange activity associated with Golgi membranes is sensitive to protease and the fungal metabolite brefeldin A^{20,21}. It remains to be seen if the reactions catalysed by ARF and Sar1 are conserved.

Several SEC12 and SARI homologues have been identified in diverse species¹², suggesting that the mechanism for vesicle budding from the ER and perhaps other compartments is accomplished by a conserved mechanism. Establishing a cycle of GTP hydrolysis and nucleotide exchange in the context of a reconstituted transport vesicle reaction provides a unique opportunity to study the role of small GTPases in membrane dynamics. □

- Salama, N., Yeung, T. & Schekman, R. *EMBO J.* (in the press).
- Bourne, H. R., Sanders, D. A. & McCormick, F. *Nature* **349**, 117–127 (1991).
- d'Enfert, C., Gensse, M. & Gaillardin, C. *EMBO J.* **11**, 4205–4211 (1992).
- Crechet, J.-B. et al. *Science* **248**, 866–868 (1990).
- Moya, M., Roberts, D. & Novick, P. *Nature* **361**, 460–463 (1993).
- Hart, J. H., Eva, A., Evans, T., Aaronson, S. A. & Cerione, R. A. *Nature* **354**, 311–313 (1991).
- Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C. & Zyllicz, M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2874–2878 (1991).
- Mockrin, S. C. & Korn, E. D. *Biochemistry* **19**, 5359–5362 (1980).
- Donaldson, J. G., Kahn, R. A., Lippincott-Schwartz, J. & Klausner, R. D. *Science* **254**, 1197–1199 (1991).
- Donaldson, J. G., Finazzi, D. & Klausner, R. D. *Nature* **360**, 350–352 (1992).
- Helms, J. B. & Rothman, J. E. *Nature* **360**, 352–354 (1992).
- Serafini, T. et al. *Cell* **67**, 239–253 (1991).
- Wuestehube, L. J. & Schekman, R. *Meth. Enzym.* **219**, 212–236 (1992).
- Sikorski, R. S. & Hieter, P. *Genetics* **122**, 19–27 (1989).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).
- Towbin, M., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4353 (1979).
- Wagner, P. et al. *EMBO J.* **6**, 2373–2379 (1987).
- Weiss, O., Jolden, J., Rulka, C. & Kahn, R. A. *J. biol. Chem.* **264**, 21066–21072 (1989).
- Jones, S., Vignais, M.-L. & Broach, J. R. *Molec. cell. Biol.* **11**, 2641–2646 (1991).

ACKNOWLEDGEMENTS. We thank J. Broach for providing a Ras2 expression construct, S. Powers for a Cdc25 expression construct, J. Donaldson for purified ARF protein, T. Yoshihisa for purified Sec23, H. Lin for technical assistance, and J. Brodsky, J. Campbell, D. Feldheim and D. Hosobuchi for comments on this manuscript. This work was supported by the Howard Hughes Medical Institute and a fellowship from the Damon Runyon-Walter Winchell Cancer Fund (C.B.).

Expression of transcription factor E2F1 induces quiescent cells to enter S phase

David G. Johnson, James K. Schwarz, W. Douglas Cress & Joseph R. Nevins

Section of Genetics, Howard Hughes Medical Institute, PO Box 3054, Duke University Medical Center, Durham, North Carolina 27710, USA

SEVERAL lines of evidence implicate the E2F transcription factor as an important component of cell proliferation control. First, E2F binding sites are found in the promoters of genes responsive to proliferation signals and the level of E2F binding activity increases at a time when many of these genes are activated^{1–3}. Second, the tumour suppressor protein Rb, as well as the related p107 protein, complexes with E2F^{2–7}, resulting in an inhibition of E2F transcriptional activity^{3,8–12}. Third, oncogenic products of the DNA tumour viruses can dissociate these E2F complexes^{4,13}. We provide here direct evidence that E2F is involved in cellular proliferation control. Specifically, we demonstrate that overexpression of the E2F1 complementary DNA^{14,15} can activate DNA synthesis in cells that would otherwise growth-arrest, with an efficiency that is similar to that achieved by the expression of the adenovirus E1A gene. Moreover, microinjection of the E2F1 cDNA into quiescent cells can induce S-phase entry, whereas two E2F1 mutants, which are unable to transactivate the DHFR and TK promoters, are unable to induce S phase. We conclude that the E2F transcription factor plays an important role in progression into S phase and that this probably coincides with its capacity to stimulate transcription.

The synthesis of DNA in REF-52 cells was measured by BrdU incorporation in cells that were transfected with various plasmids and then deprived of serum. Transfection of a plasmid expressing a β -galactosidase gene (RSV β -gal) helped the identification of cells that were successfully transfected. After transfection, cells were washed and incubated in media containing 0.1% serum for 48 h. BrdU was then added to the media and cells were incubated for an additional 30 h before fixing. Transfected cells were identified by staining for β -galactosidase and cells that entered S phase were identified by staining for BrdU incorporation (Fig. 1b). The addition of serum to a culture of quiescent REF-52 cells resulted in about 47% of the β -gal-positive cells becoming BrdU positive (Table 1). Using this assay, we measured the ability of

Received 24 May; accepted 8 July 1993.

- Nakano, A., Brada, D. & Schekman, R. *J. Cell Biol.* **107**, 851–863 (1988).
- Rexach, M. F. & Schekman, R. W. *J. Cell Biol.* **114**, 219–229 (1990).
- Kaiser, C. & Schekman, R. *Cell* **61**, 723–733 (1990).
- Nakano, A. & Muramatsu, M. *J. Cell Biol.* **109**, 2677–2691 (1989).
- Barlowe, C., d'Enfert, C. & Schekman, R. *J. Cell Biol.* **268**, 873–879 (1993).
- d'Enfert, C., Wuestehube, L. J., Lila, T. & Schekman, R. *J. Cell Biol.* **114**, 663–670 (1991).
- d'Enfert, C., Barlowe, C., Nishikawa, S., Nakano, A. & Schekman, R. *Molec. cell. Biol.* **11**, 5727–5734 (1991).
- Yoshihisa, T., Barlowe, C. & Schekman, R. *Science* **259**, 1466–1468 (1993).
- Pryer, N. K., Wuestehube, L. & Schekman, R. A. *Rev. Biochem.* **61**, 471–516 (1992).

a transfected E1A gene to stimulate BrdU incorporation. Transfection of an E1A_{12S}-expressing plasmid into quiescent REF-52 cells stimulated a large fraction (37%) of the β -gal-positive cells to become BrdU positive (Fig. 1, Table 1), consistent with previous reports¹⁶⁻²¹. Moreover, the efficiency was close to that of serum addition in the population of β -gal-positive cells.

Given the ability of E1A to stimulate DNA synthesis, together with the fact that previous experiments have shown that this activity coincides with the ability of E1A to interact with cellular proteins, many of which regulate E2F¹⁷⁻¹⁹, we tested whether deregulated expression of E2F1 can alter cell growth control. A plasmid was constructed that placed the E2F1 cDNA under the control of the CMV enhancer/promoter. Two E2F1 mutants, one which lacks the normal C-terminal acidic activation domain and one which has a mutation in the DNA binding domain, were also created (Fig. 1a). Transfection of the E2F1 cDNA resulted in a substantial fraction of the transfected cells becoming BrdU positive (Fig. 1c, Table 1). The number of positive cells ranged from a low of 24% of the transfected cells to as many as 45%. In contrast, transfection of either of the E2F1 mutants resulted in no increase over background in the number of BrdU-positive cells. We therefore conclude that the over-

expression of E2F1 can block entry into quiescence in serum-starved cells.

These E2F1 constructs were also tested for their ability to activate transcription in serum-starved REF-52 cells. Each E2F1-expressing plasmid was cotransfected with a CAT gene under the control of the DHFR or TK promoter. E2F binding sites can be found in the promoters of these genes, in regions shown to be important for S-phase induction of transcription²²⁻²⁵. In fact, mutation of the E2F sites in the DHFR promoter abolishes the cell cycle regulation of this gene²⁴. Cotransfection of the E2F1 cDNA stimulated the DHFR and TK gene promoters 6- and 20-fold, respectively (Fig. 2a). In contrast, the E2F1 mutants had only a modest effect on either promoter. It is thus clear that the cloned E2F1 product can stimulate transcription of two S-phase regulated genes in the absence of normal growth stimulatory signals.

The experiments summarized in Table 1 clearly demonstrate that overexpression of E2F1 can alter cell growth control. Nevertheless, it is not possible to determine from this experiment whether E2F1 expression blocks cell cycle exit and thus maintains cells in a proliferative state or whether E2F1 expression can drive an otherwise quiescent cell to enter S phase. To address

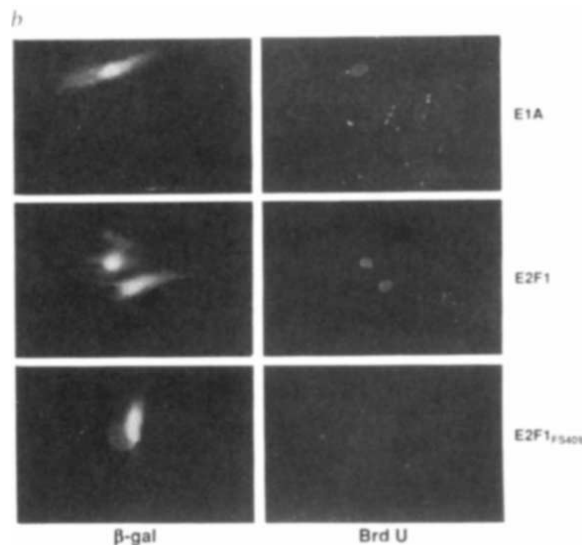
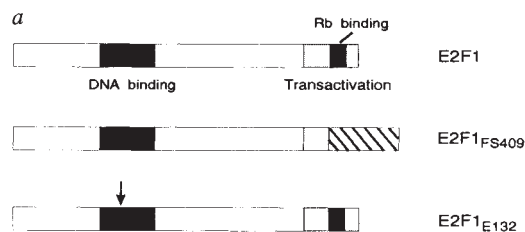


FIG. 1 Induction of DNA synthesis in serum-starved cells. a, Schematic representation of the E2F1 constructs. The functional domains of the E2F1 cDNA-encoded protein are depicted at the top. Shown below is the comparison of the wild-type E2F1 cDNA with the E2F1 mutants. The E2F1 cDNA was cloned into the vector pcDNA1/Amp (Invitrogen) as a *Bam*HI-*Xba*I fragment. To create the E2F1_{FS409} mutant, a nucleotide was deleted at the codon specifying amino-acid 409 creating a frameshift mutation at that point. The resulting protein is thus lacking the C-terminal 28 amino acids of E2F1 and is 63 amino acids longer than the wild-type protein. This substitution removes most of the previously defined transcriptional activating domain^{14,15}. To create the E2F1_{E132} mutant, the DNA sequence at E2F1 codons L132 and N133 (CTGAAT) was replaced with an *Eco*RI restriction site sequence (GAATCC). b, REF-52 cells were transfected with RSV- β -gal together with CMV-E1A_{12S}, CMV-E2F1, or CMV-E2F1_{FS409}. Cells were then visualized for β -galactosidase expression (left panels) or BrdU incorporation (right panels). c, Quantification of BrdU-positive REF-52 cells. The data plotted is the average of the experiments presented in Table 1 with the exception of the E2F1_{E132} mutant that represents a single experiment.

METHODS. REF-52 cells were transfected by the calcium phosphate coprecipitation procedure using 5 μ g RSV β -gal, pGEM4 as a nonspecific carrier, and 200 ng CMV E1A_{12S} or 500 ng CMV-E2F1, where appropriate, for a total of 20 μ g in 1 ml. The DNA precipitate (0.2 ml) was added to each 35 mm dish containing a cover slip. Cells were exposed to the DNA precipitates for 16 h, washed, and then placed in α -MEM containing 0.1% fetal calf serum for 48 h. BrdU (final concentration of 10 μ M) was then added and the cells were incubated for another 30 h before fixing. Cells were fixed with 4% paraformaldehyde in PBS for 10 min at room temperature then dehydrated with methanol/acetone (1:1) for 30 s at room temperature. Staining for β -galactosidase used a rabbit polyclonal antibody (1:500 in 1% BSA/PBS) (5 Prime, 3 Prime) and FITC-conjugated goat anti-rabbit IgG (Sigma). Coverslips were fixed again in 4% paraformaldehyde and then were treated with 2M HCl for 1 h at 37 $^{\circ}$ C and subsequently stained with an α -BrdU monoclonal antibody as suggested by the supplier (Boehringer-Mannheim). Texas red-conjugated goat anti-mouse IgG (Sigma) was used as a secondary antibody to visualize BrdU staining. Coverslips were then examined using fluorescence microscopy with the appropriate filters.

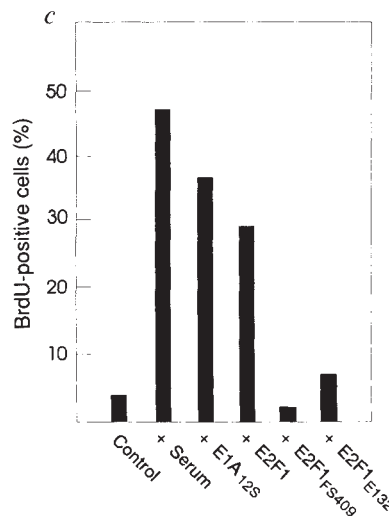


TABLE 1 Effect of E1A and E2F1 expression on BrdU incorporation in transfected REF-52 cells

Experiment	Treatment	Number of β -gal-positive cells	Number of BrdU-positive cells	%BrdU positive
1	Control	69	5	7
	+ Serum	102	47	46
	+ E1A	60	23	38
	+ E2F1	40	16	40
2	Control	87	1	1
	+ Serum	172	82	48
	+ E1A	89	38	43
	+ E2F1	114	27	24
	+ E2F1 _{FS409}	93	4	4
3	Control	406	7	2
	+ Serum	434	204	47
	+ E1A	68	19	28
	+ E2F1	260	62	24
	+ E2F1 _{FS409}	321	3	1
4	Control	158	15	9
	+ Serum	226	106	47
	+ E2F1	92	41	45
	+ E2F1 _{E132}	88	6	7

REF-52 cells were transfected and scored for β -gal expression as well as BrdU incorporation as described in the legend to Fig. 1.

TABLE 2 BrdU incorporation after microinjection of quiescent REF-52 cells with E2F1

Experiment	Treatment	Number of β -gal-positive cells	Number of BrdU-positive cells	%BrdU positive
1	Control	97	1	1
	E2F1	83	20	24
	E2F1 _{FS409}	67	0	0
	E2F1 _{E132}	81	4	5
2	Control	29	0	0
	E2F1	29	4	13
	E2F1 _{FS409}	25	1	4
	E2F1 _{E132}	24	0	0
3	Control	39	0	0
	E2F1	61	9	15
	E2F1 _{FS409}	31	0	0
	E2F1 _{E132}	43	0	0
4	Control	33	0	0
	E2F1	36	5	14
	E2F1 _{FS409}	21	0	0
	E2F1 _{E132}	24	0	0

REF-52 cells were brought to quiescence by serum starvation and then microinjected with the indicated plasmids. The experimental details are described in the legend to Fig. 2b.

this question, we microinjected the E2F1 cDNA into REF-52 cells that were brought to quiescence by serum starvation. After microinjection, DNA synthesis was again scored by measuring BrdU incorporation. The E2F1 cDNA was indeed capable of inducing S phase in otherwise quiescent cells (Fig. 2b, Table 2). In contrast, the two E2F1 mutants, one lacking the transcription activation domain and the other bearing a mutation in the DNA-binding domain, were unable to stimulate BrdU incorporation. Because the product of the DNA-binding domain mutant retains the ability to interact with the Rb protein as well as the p107 polypeptide³¹, it would appear that the capacity of the wild-type

E2F1 cDNA product to promote DNA synthesis is not a simple consequence of the titration of Rb or p107.

Previous experiments have shown that the Rb protein blocks cell growth in G1 and that this depends on Rb sequences that are critical for E2F interaction^{8,26-28}. This suggests that Rb-mediated inhibition of E2F transcriptional activity may, at least in part, be responsible for the ability of Rb to inhibit progression into S phase; if true, our data suggests that the overexpression of E2F1 overcomes this Rb block and allows the otherwise quiescent cells to enter S phase. Although the precise mechanism for the E2F1-mediated S-phase induction remains to be determined,

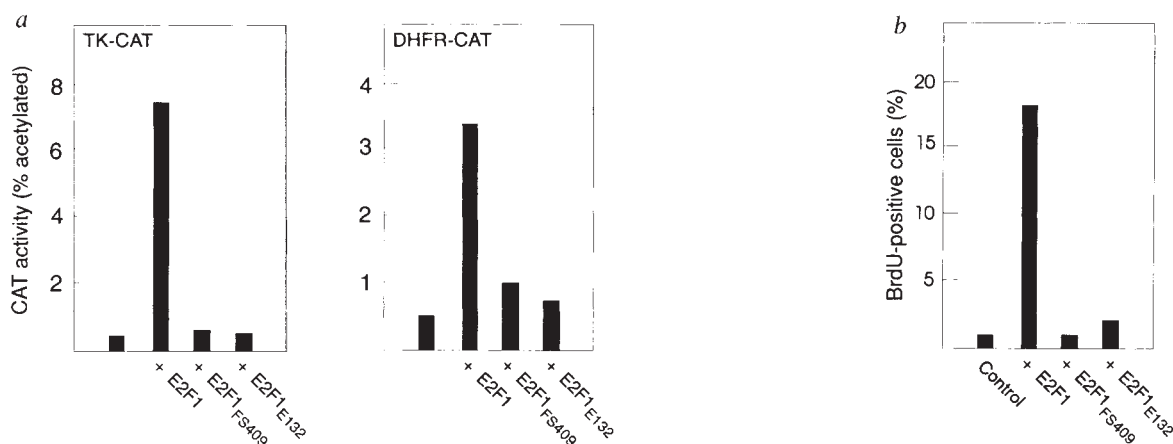


FIG. 2 a, Expression of E2F1 activates transcription in serum starved cells. REF52 cells were split 1:5 in DMEM containing 10% fetal calf serum 24 h before transfection. Calcium phosphate precipitates were formed using 10 μ g target CAT plasmid, 5 μ g RSV β -gal, salmon sperm DNA (Specialty Media, Inc.) as a nonspecific carrier DNA, and 300 ng CMV E2F1 or the CMV-E2F1 mutants. The DHFR-CAT construct has been described previously²⁵. The TK-CAT construct is derived from the human thymidine kinase gene and contains 123 nucleotides upstream and 37 nucleotides downstream of the major transcription start site²². The total DNA concentration was 20 μ g in 1 ml. The DNA precipitate was exposed to the cells for 16 h, cells were washed, and placed in DMEM containing 0.5% FCS. Cells were collected 48 h later and cell

lysates were prepared. CAT assays were as described previously⁸ using 75% of the cell lysates and incubating the reactions for 12 h. β -galactosidase activity was also assayed as described previously⁸ and used as an internal control to correct for transfection efficiency. b, E2F1 induces DNA synthesis in quiescent REF-52 cells. Cultures of REF-52 cells were brought to quiescence by growth in DMEM containing 0.1% fetal calf serum for 40–48 h. Cells were then microinjected with the indicated plasmids (DNA at a concentration of 8 ng ml⁻¹). After microinjection, BrdU was added to the medium and incubation continued for 30 h. Details of staining for β -galactosidase and BrdU were as described in Fig. 1. The data plotted is the average of the experiments in Table 2.

and several possible mechanisms can be envisaged, we favour the idea that accumulation of uncomplexed E2F, as normally occurs in late G1^{2,3}, could be responsible for the S-phase induction. This would be consistent with the ability of E2F1 to activate cellular genes, such as DHFR and thymidine kinase, that are necessary for DNA synthesis. Moreover, recent experiments have shown that the E2F1 gene is induced in late G1, dependent on new protein synthesis, and probably represents a target for events occurring earlier in G1^{29,30}. It is thus possible that overexpression of E2F1 in a GO cell could bypass the need for many G1 events. Regardless of the precise mechanism, the results clearly show that E2F1 is an important participant in the process of cell-cycle progression. □

Received 7 April; accepted 2 July 1993.

1. Nevins, J. R. *Science* **258**, 424–429 (1992).
2. Mudryj, M. et al. *Cell* **65**, 1243–1253 (1991).
3. Schwarz, J. K. et al. *EMBO J.* **12**, 1013–1020 (1993).
4. Chellappan, S. P., Hiebert, S., Mudryj, M., Horowitz, J. M. & Nevins, J. R. *Cell* **65**, 1053–1061 (1991).
5. Bagchi, S., Weinmann, R. & Raychaudhuri, P. *Cell* **54**, 1063–1072 (1991).
6. Chittenden, T., Livingston, D. M. & Kaelin, W. G. *Cell* **65**, 1073–1082 (1991).
7. Bandara, L. R. & LaThangue, N. B. *Nature* **351**, 494–497 (1991).
8. Hiebert, S. W., Chellappan, S. P., Horowitz, J. M. & Nevins, J. R. *Genes Dev.* **6**, 177–185 (1992).
9. Zamanian, M. & LaThangue, N. B. *EMBO J.* **11**, 2603–2610 (1992).
10. Weintraub, S. J., Prater, C. A. & Dean, D. C. *Nature* **358**, 259–261 (1992).

11. Hamel, P. A., Gill, R. M., Phillips, R. A. & Gallie, B. L. *Molec. cell. Biol.* **12**, 3431–3438 (1992).
12. Dalton, S. *EMBO J.* **11**, 1797–1804 (1992).
13. Chellappan, S. P. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 4549–4553 (1992).
14. Helin, K. et al. *Cell* **70**, 337–350 (1992).
15. Kaelin, W. G. et al. *Cell* **70**, 351–364 (1992).
16. Stabel, S., Argos, P. & Philipson, L. *EMBO J.* **4**, 2329–2336 (1985).
17. Lillie, J. W., Loewenstein, P. M., Green, M. R. & Green, M. *Cell* **50**, 1091–1100 (1987).
18. Moran, B. & Zierler, B. *Molec. cell. Biol.* **8**, 1756–1764 (1988).
19. Howe, J. A., Mymyk, J. S., Egan, C., Branton, P. E. & Bayley, S. T. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5883–5887 (1990).
20. Kaczmarek, L., Ferguson, B., Rosenberg, M. & Baserga, R. *Virology* **152**, 1–10 (1986).
21. Spindler, K. R., Eng, C. Y. & Berk, A. J. *J. Virol.* **53**, 742–750 (1985).
22. Lipson, K. E. & Baserga, R. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9774–9777 (1989).
23. Roehl, H. H. & Conrad, S. E. *Molec. cell. Biol.* **10**, 3834–3837 (1990).
24. Means, A. L., Slansky, J. E., McMahon, S. L., Knuth, M. W. & Farnham, P. J. *Molec. cell. Biol.* **12**, 1054–1063 (1992).
25. Blake, M. C. & Azizkhan, J. C. *Molec. cell. Biol.* **9**, 4994–5002 (1989).
26. Goodrich, D. W., Wang, N. P., Qian, Y.-W., Lee, E. Y.-H. P. & Lee, W.-H. *Cell* **67**, 293–302 (1991).
27. Qin, X.-Q., Chittenden, T., Livingston, D. M. & Kaelin, W. G. Jr. *Genes Dev.* **6**, 953–964 (1992).
28. Qian, Y., Luckey, C., Horton, L., Esser, M. & Templeton, D. J. *Molec. cell. Biol.* **12**, 5363–5372 (1992).
29. Slansky, J. E., Li, Y., Kaelin, W. G. & Farnham, P. J. *Molec. cell. Biol.* **13**, 1610–1618 (1993).
30. Shirodkar, S. et al. *Cell* **68**, 157–166 (1992).
31. Cross, W. D., Johnson, D. J. & Nevins, J. R. *Molec. cell. Biol.* (in the press).

ACKNOWLEDGEMENTS. D. G. J. and J. K. S. made equal contributions to this work. We thank E. Harlow and B. Kaelin for the gifts of the E2F1 cDNAs, K. Burridge for the REF-52 cells, L. Babiss and M. Bonham for help with the microinjection experiments and K. Culler and T. Jenkins for help in the preparation of the manuscript. D. G. J. is the recipient of an American Cancer Society postdoctoral fellowship, J. K. S. is supported by a Pfizer Physicians postdoctoral fellowship, and W. D. C. is the recipient of a Damon Runyon-Walter Winchell postdoctoral fellowship.

The T-cell transcription factor NFAT_p is a substrate for calcineurin and interacts with Fos and Jun

Jugnu Jain*, Patricia G. McCaffrey*, Zoe Miner*†, Tom K. Kerppola†, John N. Lambert§, Gregory L. Verdine§, Tom Curran† & Anjana Rao*||

* Division of Tumor Virology, Dana-Farber Cancer Institute, and Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115, USA

† Department of Molecular Oncology and Virology, Roche Institute of Molecular Biology, Roche Research Center, Nutley, New Jersey 07110, USA

§ Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138, USA

TRANSCRIPTION of lymphokine genes in activated T cells is inhibited by the immunosuppressive agents cyclosporin A and FK506, which act by blocking the phosphatase activity of calcineurin^{1–3}. NFAT, a DNA-binding protein required for interleukin-2 gene transcription, is a potential target for calcineurin, cyclosporin A and FK506^{4–11}. NFAT contains a subunit (NFAT_p) which is present in unstimulated T cells and which forms a complex with Fos and Jun proteins in the nucleus of activated T cells^{9,11}. Here we report that NFAT_p is a DNA-binding phosphoprotein of relative molecular mass ~120,000 and is a substrate for calcineurin *in vitro*. Purified NFAT_p forms DNA–protein complexes with recombinant Jun homodimers or Jun–Fos heterodimers; the DNA-binding domains of Fos and Jun are essential for the formation of the NFAT_p–Fos–Jun–DNA complex. The interaction between the lymphoid-specific factor NFAT_p and the ubiquitous transcription factors Fos and Jun provides a novel mechanism for combinatorial regulation of interleukin-2 gene transcription, which integrates the

calcium-dependent and the protein-kinase C-dependent pathways of T-cell activation.

NFAT_p purified from a murine T-cell line migrates on SDS–polyacrylamide gels as a diffuse band with an apparent *M_r* of 120,000 (~120K) (Fig. 1a, lane 1). Treatment of this protein with calcineurin resulted in a significant decrease in its apparent *M_r* and the resolution of four sharply defined bands grouped into two doublets (Fig. 1a, lane 2). Both changes were due to dephosphorylation rather than proteolysis, as they were blocked by the peptide inhibitor of calcineurin¹² and by the general phosphatase inhibitor pyrophosphate (data not shown). We conclude that NFAT_p consists of one or more ~120K phosphoproteins that can be dephosphorylated by calcineurin *in vitro*. The four sharp bands generated by phosphatase treatment may represent spliced variants of a single gene, differentially modified forms of a single polypeptide, or distinct polypeptides capable of binding to the NFAT site.

Calcineurin-mediated dephosphorylation of NFAT_p affected neither its ability to bind to the NFAT site of the murine interleukin-2 (IL-2) promoter nor its ability to form complexes with Fos and Jun proteins⁹. Both untreated and calcineurin-treated NFAT_p formed a single protein–DNA complex (Fig. 1b, lanes 1, 3) and both formed an additional DNA–protein complex with recombinant Fos and Jun¹³ (lanes 2, 4, see below for specificity controls). These data are consistent with the hypothesis¹ that dephosphorylation of NFAT_p by calcineurin in activated T cells may regulate its subcellular localization rather than its ability to bind to DNA or associate with Fos and Jun.

The reconstituted complex formed by NFAT_p, Fos and Jun (Fig. 2a, lane 2, upper arrowhead) had the same DNA-binding specificity as the protein–DNA complex formed with NFAT_p alone (Fig. 2a, lane 1), based on competition with mutated NFAT oligonucleotides⁹. Formation of both complexes was inhibited by excess unlabelled NFAT oligonucleotide and by the M1 and M2 but not the M3 mutant NFAT oligonucleotides (Fig. 2a, compare lane 2 with lanes 3–6). In contrast, formation of the upper complex was selectively eliminated by competition with an excess of unlabelled AP-1 oligonucleotide (Fig. 2a, lane 7) or by antibodies to Fos and Jun peptides^{9,14} (data not shown), indicating that only the upper complex contained Fos and Jun. In the presence of Ref-1, the redox factor which stimulates the DNA-binding activity of Fos and Jun^{15,16}, all of the NFAT_p

† Present address: Centre for Nutrition and Medicine, New England Deaconess Hospital, 1 Autumn Street, Boston, Massachusetts 02115, USA.

|| To whom correspondence should be addressed at the Dana-Farber Cancer Institute.

FIG. 1 Purified NFAT_p is a ~120K phosphoprotein and a substrate for calcineurin (CaN). **a**, Migration of phosphorylated and dephosphorylated NFAT_p on SDS-polyacrylamide gels. Purified NFAT_p (lane 1) treated with CaN and calmodulin (lane 2) was separated on a 6% polyacrylamide-SDS gel and then silver stained (Pierce Gelcode kit). Lane 3 shows CaN and calmodulin alone. Purified NFAT_p, either untreated or treated with calcineurin, was tested for binding to the NFAT oligonucleotide in the absence (lanes 1, 3) or presence (lanes 2, 4) of c-Fos(T) and c-Jun(T) (300 nM each) in an electrophoretic mobility shift assay (EMSA)¹¹. **METHODS.** NFAT_p was purified from cytosolic extracts of a *v-fos*-transformed murine T-cell clone²⁸ by successive heparin-agarose and DNA-affinity chromatography (P.G.McC. *et al.*, manuscript in preparation). The purified protein was mixed with calcineurin (Sigma, 150 nM) and calmodulin (Sigma, 330 nM) in a buffer containing 5 mM HEPES pH 7.4, 25 mM NaCl, 15 mM β -mercaptoethanol, 1.5 mM MnCl₂, 1 mM EDTA, 2.5% glycerol and 200 μ M inhibitory peptide¹² or 30 mM sodium pyrophosphate. After incubation for 10 min at 30 °C, the samples were boiled for 10 min in Laemmli sample buffer and fractionated on 6% acrylamide-SDS gels. Reconstitution with recombinant Jun(T)(199–334) and Fos(T)(139–200) was as described in Fig. 2 legend.

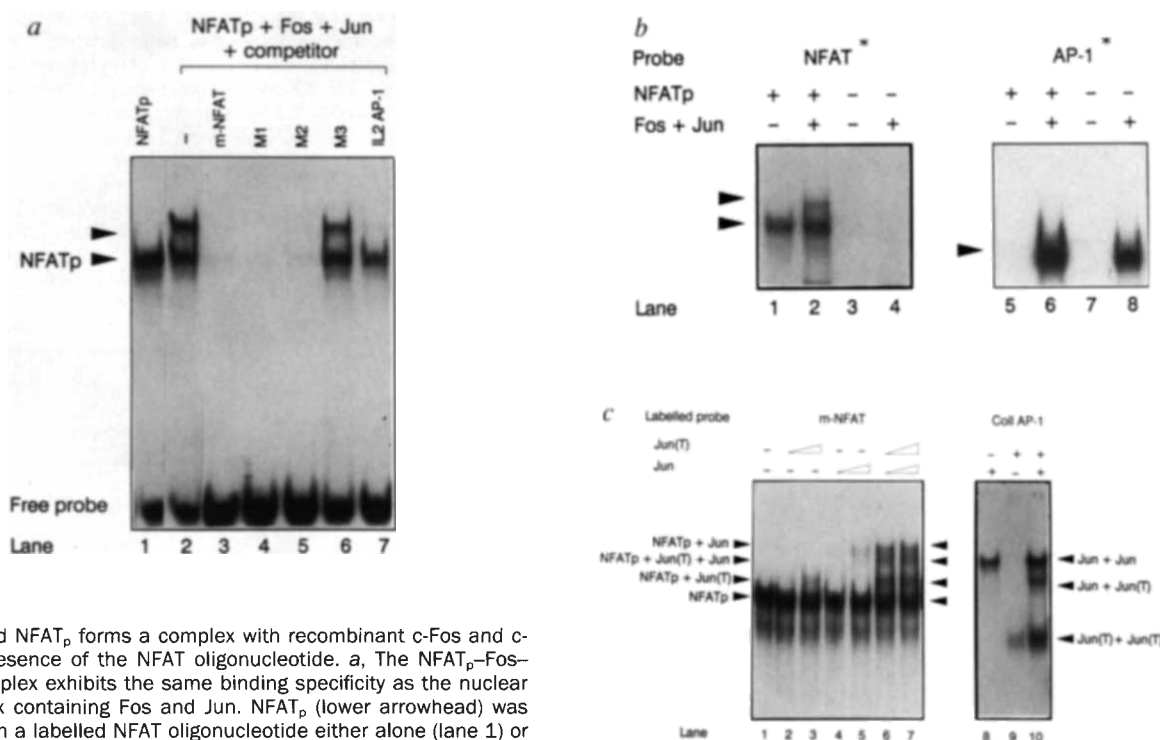
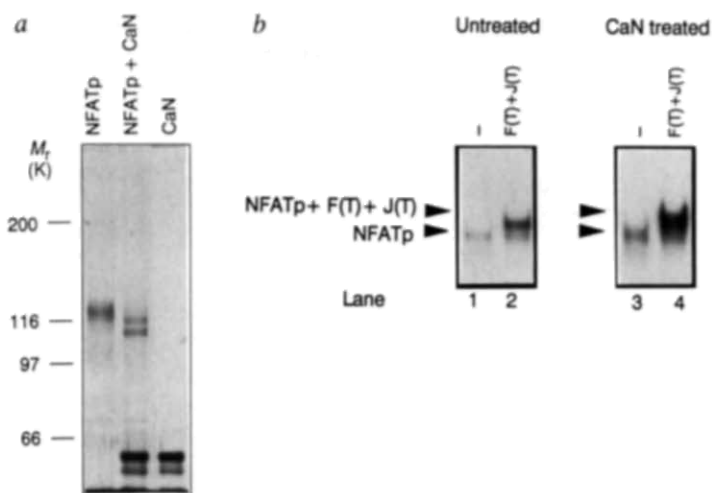


FIG. 2 Purified NFAT_p forms a complex with recombinant c-Fos and c-Jun in the presence of the NFAT oligonucleotide. **a**, The NFAT_p-Fos-Jun-DNA complex exhibits the same binding specificity as the nuclear NFAT complex containing Fos and Jun. NFAT_p (lower arrowhead) was incubated with a labelled NFAT oligonucleotide either alone (lane 1) or in the presence of recombinant c-Fos and c-Jun proteins (lanes 2–7). The upper arrowhead indicates the NFAT_p-Fos-Jun complex. For lanes 3–7, the binding reactions contained a 200-fold excess of unlabelled competitor oligonucleotides as indicated. **b**, Fos and Jun do not bind to the NFAT oligonucleotide in the absence of NFAT_p. Purified c-Fos and c-Jun were incubated with the labelled NFAT oligonucleotide in the presence (lanes 1, 2) or absence (lanes 3, 4) of NFAT_p. The same concentrations of Fos and Jun were also tested for binding to the AP-1 site of the collagenase promoter in the presence (lanes 5, 6) or absence (lanes 7, 8) of NFAT_p. The NFAT_p complex (left lower arrowhead) and the NFAT_p-Fos-Jun complex (left upper arrowhead) bound to the NFAT oligonucleotide, and the single Fos-Jun complex (right panel's arrowhead) bound to the AP-1 oligonucleotide, are indicated by arrows. **c**, The NFAT_p-Jun complex contains a Jun homodimer. NFAT_p (lane 1) was incubated with the labelled NFAT oligonucleotide and Jun(T) (residues 199–334) alone (lanes 2, 3; 70 and 140 nM, respectively), full length c-Jun alone (lanes 4, 5; 100 and 200 nM, respectively) or both

(lanes 6, 7). The right panel shows the binding of full-length c-Jun alone (lane 8), Jun(T) alone (lane 9), or both full-length and Jun(T) proteins (lane 10) to the labelled collagenase AP-1 oligonucleotide under the same conditions. The DNA-protein complexes observed are indicated by arrows.

METHODS. Full-length and truncated c-Fos and c-Jun recombinant proteins were expressed in *Escherichia coli* as hexahistidine fusion proteins and purified to homogeneity from *E. coli* lysates as described^{13,19}. Binding reactions (15 μ l) contained 2–3 units (~150 pg) of NFAT_p (diluted in 5 mg ml⁻¹ BSA), 17 μ g ml⁻¹ dIdC, 6 μ l binding buffer (20 mM HEPES pH 7.5, 250 mM NaCl, 20% glycerol, 0.5 mM dithiothreitol (DTT), 50 mM c-Fos and/or c-Jun proteins (diluted in 25 mM HEPES, pH 7.5, 5% glycerol, 1 mM DTT) and ~10,000 c.p.m. (0.2–0.5 ng) of labelled oligonucleotides (sequences in ref. 9). Binding reactions with the AP-1 oligonucleotide contained 5 mM DTT.

present in the binding reaction could be recruited into an NFAT_p-Fos-Jun-DNA complex (data not shown), indicating that the diffusely migrating NFAT_p band of Fig. 1a, was homogeneous with respect to its ability to associate with Fos and Jun even though it appeared heterogeneous after calcineurin treatment.

In the absence of NFAT_p, Fos and Jun did not bind to the

labelled NFAT oligonucleotide (Fig. 2b, compare lanes 2 and 4). Likewise, the NFAT oligonucleotide did not compete even at high concentrations for binding of Fos-Jun dimers to AP-1 site oligonucleotides (data not shown). Binding of Fos and Jun to an AP-1 site, however, was not significantly altered in the presence or absence of NFAT_p (Fig. 2b, compare lanes 6 and 8). Thus, the formation of the NFAT_p-Fos-Jun-DNA complex

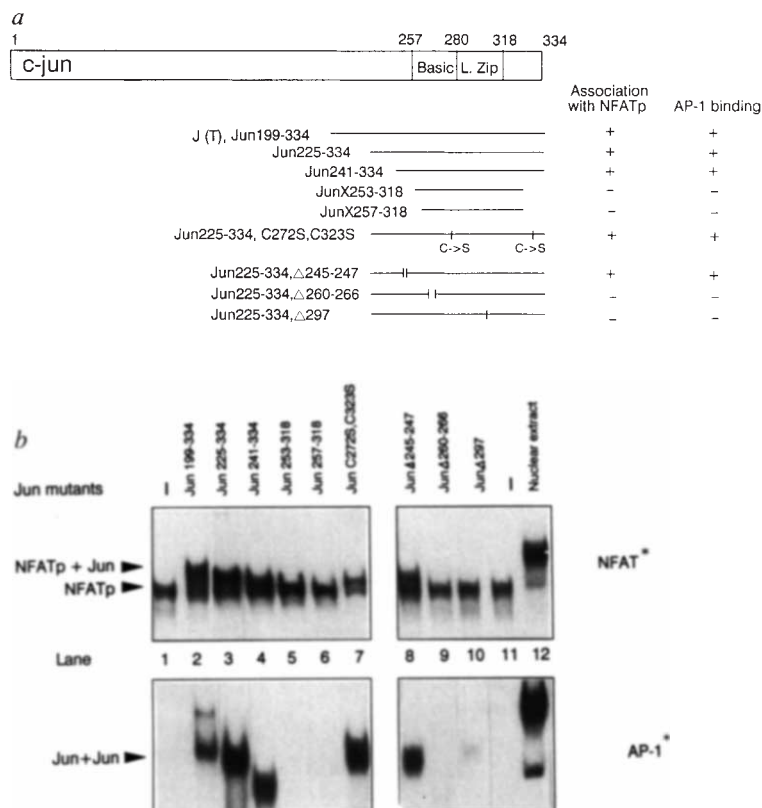
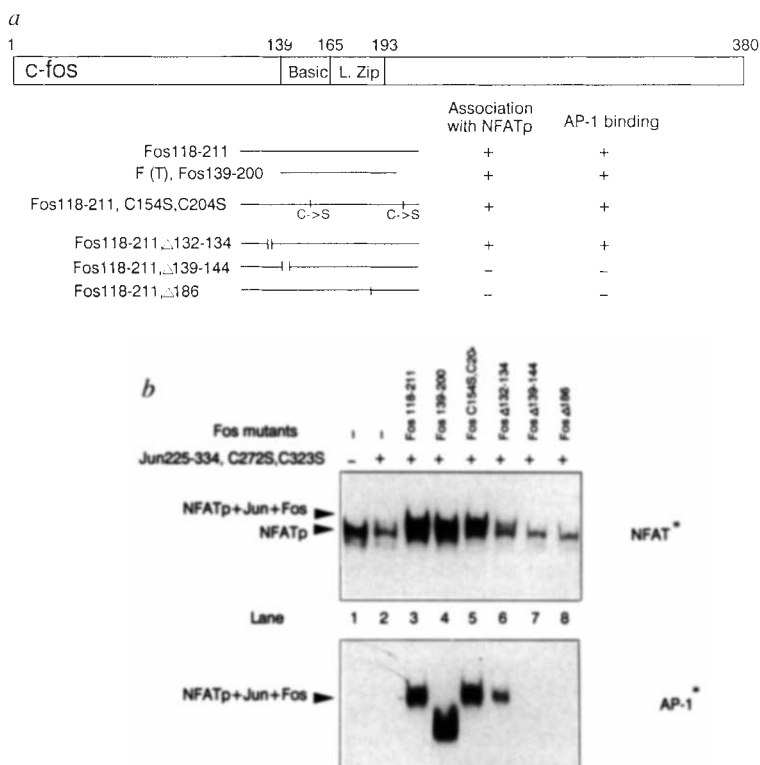


FIG. 3 The ability of truncated and mutant c-Jun proteins to form an NFAT complex correlates precisely with their ability to bind to the collagenase AP-1 site. **a**, Diagram of the c-Jun proteins tested, scored for their ability to bind to the collagenase AP-1 site and to associate with NFAT_p. **b**, Purified NFAT_p (lanes 1, 11) was incubated with the labelled NFAT oligonucleotide (upper panel) or the labelled collagenase AP-1 oligonucleotide (lower panel) in the presence of the truncated or mutant c-Jun (200–300 nM) shown schematically in **a** (lanes 3–10). Formation of DNA–protein complexes was monitored by EMSA^{9,14}. The truncated proteins Jun(T), Jun(225–334) and Jun(241–334), the internally deleted proteins Jun(225–334, Δ245–247) and the serine substitution mutant Jun(225–334, C272S, C323S) all bind effectively to an AP-1 site (bottom panel, lanes 2–4, 7, 8); they also effectively form complexes with NFAT_p (top panel, lanes 2–4, 7, 8). In contrast, Jun(253–318), Jun(257–318) and Jun(225–334, Δ260–266), which contain deletions in the DNA-binding domain, and Jun(225–334, Δ297), in which the third leucine of the leucine zipper has been deleted, bind poorly to the AP-1 site and do not form a complex with NFAT_p (lanes 5, 6, 9). Lane 12 shows the binding of nuclear extracts from stimulated Ar-5 T cells⁹ to the labelled oligonucleotides. Lane 2, Jun(T) Jun199–334; lane 3, Jun225–334; lane 4, Jun241–334; lane 5, Jun253–318; lane 6, Jun257–318; lane 7, Jun225–334, C272S, C323S; lane 8, Jun225–334, Δ245–247; lane 9, Jun225–334, Δ260–266; lane 10, Jun225–334, Δ297. Jun253–318 and Jun257–318 proteins contain the factor X cleavage site inserted between the hexahistidine fusion peptide and the Jun-coding region.

Fig. 4 The ability of truncated and mutant Fos proteins to form NFAT_p-Jun-Fos complexes correlates precisely with their ability to bind as the corresponding Jun-Fos dimers to the collagenase AP-1 site. **a**, Diagram of the c-Fos proteins tested, scored for their ability to bind as Fos-Jun dimers to the collagenase AP-1 site and to associate with NFAT_p. **b**, Purified NFAT_p (lane 1) was incubated with the labelled NFAT oligonucleotide (upper panel) or the labelled collagenase AP-1 oligonucleotide (lower panel) in the presence of Jun(225–334, C272S, C323S) (lane 2) and the truncated or mutant c-Fos proteins (200–300 nM) shown schematically in (lanes 3–8). The Jun protein was used at a low concentration so that the binding of Jun homodimers to the AP-1 site or their association with NFAT_p was not detected. The truncated proteins Fos(T) and Fos(118–211), the internally deleted protein Fos(118–211, Δ132–134) and the serine substituted protein Fos(118–211, C154S, C204S) that could bind to the AP-1 site (bottom panel, lanes 3–6) could form a complex with NFAT_p (top panel, lanes 3–6). In contrast, the basic region deletion mutant Fos(Fos118–211, Δ139–144) and the leucine zipper mutant Fos(118–211, Δ186) neither bound to the AP-1 site nor formed a complex with NFAT_p (lanes 5, 6). Lane 3, Fos 118–211; lane 4, Fos(T) Fos139–200; lane 5, Fos118–211, C154S, C204S; lane 6, Fos118–211, Δ132–134; lane 7, Fos118–211, Δ139–144; lane 8, Fos118–211, Δ186.



requires the presence of DNA containing a specific NFAT-binding site, and does not occur in the presence of AP-1 site DNA. It is possible that Fos and Jun contact an AP-1-like sequence (TGTTTCA) in the NFAT oligonucleotide; an oligonucleotide mutated in this sequence¹⁷ still binds NFAT_p but does not permit formation of the NFAT_p-Fos-June-DNA complex (data not shown).

In the absence of Fos, Jun proteins could form specific complexes with NFAT_p and labelled NFAT oligonucleotide (Fig. 2c). A mixture of full-length and truncated Jun was used to show that Jun associates with NFAT_p as a homodimer. When bound to the collagenase AP-1 oligonucleotide, individual full-length (Jun) and truncated (Jun(T)¹⁸) protein formed single homodimeric complexes as expected (Fig. 2a, lanes 8, 9), whereas the mixture of both proteins formed three distinguishable complexes (lane 10). Similarly, either full-length or truncated Jun formed a single species of NFAT_p-Jun-DNA complex (Fig. 2c, lanes 3, 5), whereas the mixture of full-length and truncated Jun formed three distinct complexes (lanes 6, 7). These results imply that the NFAT complex, like the AP-1 complex, can be formed either with Jun-Jun homodimers or with Jun-Fos heterodimers.

Analysis of a panel of truncated, internally deleted and mutant derivatives of Jun^{15,18,20} showed that the ability of Jun dimers to form a complex with NFAT_p correlated with their ability to bind specifically to the AP-1 site (Fig. 3). Similarly, the ability of truncated, internally deleted or mutant Fos proteins^{19,20} to form a complex with NFAT_p in the presence of a functional Jun protein also correlated with their ability to bind to the AP-1 site (Fig. 4). Therefore, Fos and Jun associate with the NFAT_p-DNA complex either as Jun homodimers or as Fos-Jun heterodimers, and they cannot form these complexes unless their DNA-binding regions are structurally intact. The most likely explanation is that Fos and Jun are in direct contact with DNA when they are part of the NFAT_p-Fos-Jun-DNA complex. NFAT_p may help binding of Fos and Jun to the AP-1-like sequence in the NFAT oligonucleotide either directly via a protein-protein interaction²⁰ or indirectly via a structural change in DNA²¹.

Thus, we have shown that NFAT_p is a ~120K phosphoprotein that is a direct substrate for calcineurin *in vitro*. The induction of Fos and Jun family members in T cells is dependent on protein kinase C¹⁴; hence, the formation of nuclear NFAT_p-Fos-Jun-DNA complexes⁹ represents a convergence of the dual requirement for calcium and protein kinase C in T-cell activation. Our own work (J. J. *et al.*, manuscript in preparation) and that of others^{17,22} indicates that multiple members of the Fos and Jun family are present in nuclear NFAT complexes; thus, the relevant interactions *in vivo* may also involve multiple Fos and Jun family members, and potentially other bZIP proteins as well. The mechanism of cooperation between NFAT_p and Fos-Jun proteins may vary for different lymphokine genes; in the granulocyte-macrophage colony-stimulating factor/interleukin-3 (GM-CSF/IL-3) enhancer, an NFAT site is located adjacent to a functional AP-1 site capable of independent binding to Fos and Jun²³. Members of the Fos and Jun family of proteins are involved in widespread interactions with other transcription factors, including glucocorticoid receptors^{24,25} and MyoD^{26,27}. Their interaction with NFAT_p reflects a potentially general mechanism by which ubiquitous transcription factors such as Fos and Jun are targeted to a DNA sequence that does not independently bind AP-1 proteins, by virtue of their association with a cell-type-specific DNA-binding protein such as NFAT_p. This type of interaction could represent a common mechanism for cooperative interactions among transcription factors, required for the tissue-specific regulation of gene expression. □

3. Sigal, N. H. & Dumont, F. J. *Rev. Immun.* **10**, 519-560 (1992).
4. Shaw, J.-P. *et al. Science* **241**, 202-205 (1988).
5. Mattila, P. S. *et al. EMBO J.* **9**, 4425-4433 (1990).
6. Brabletz, T., Pietrowski, I. & Serfling, E. *Nucleic Acids Res.* **19**, 61-67 (1991).
7. Banerji, S. S., Parsons, J. N. & Tocci, M. J. *Molec. cell. Biol.* **11**, 4074-4087 (1991).
8. Flanagan, W. M., Corthesy, B., Bram, R. J. & Crabtree, G. R. *Nature* **352**, 803-807 (1991).
9. Jain, J., McCaffrey, P. G., Valge-Archer, V. E. & Rao, A. *Nature* **356**, 801-804 (1992).
10. Clifton, N. A. & Crabtree, G. R. *Nature* **357**, 695-697 (1992).
11. McCaffrey, P. G., Perrino, B. A., Soderling, T. R. & Rao, A. *J. biol. Chem.* **268**, 3747-3752 (1993).
12. Hashimoto, Y., Perrino, B. A. & Soderling, T. R. *J. biol. Chem.* **265**, 1924-1927 (1990).
13. Abate, C., Luk, D., Gagne, E., Roeder, R. G. & Curran, T. *Molec. cell. Biol.* **10**, 5532-5535 (1990).
14. Jain, J., Valge-Archer, V. E. & Rao, A. *J. Immun.* **148**, 1240-1250 (1992).
15. Abate, C., Patel, L., Rauscher, F. J. & Curran, T. *Science* **249**, 1157-1161 (1990).
16. Xanthoudakis, S. & Curran, T. *EMBO J.* **11**, 653-665 (1992).
17. Boise, L. H. *et al. Molec. cell. Biol.* **13**, 1911-1919 (1993).
18. Abate, C., Luk, D. & Curran, T. *Molec. cell. Biol.* **11**, 3624-3632 (1991).
19. Abate, C., Luk, D., Gentz, R., Rauscher, F. J. & Curran, T. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1032-1036 (1990).
20. Patel, L., Abate, C. & Curran, T. *Nature* **347**, 572-575 (1990).
21. Kerppola, T. K. & Curran, T. *Cell* **66**, 317-326 (1991).
22. Northrop, J. P., Ullman, K. S. & Crabtree, G. R. *J. biol. Chem.* **268**, 2917-2923 (1993).
23. Cockerill, J. P., Shannon, M. F., Bert, A. G., Ryan, G. R. & Vadas, M. A. *Proc. natn. Acad. Sci. U.S.A.* **90**, 2466-2470 (1993).
24. Miner, J. N. & Yamamoto, K. R. *Genes Dev.* **6**, 2491-2501 (1992).
25. Kerppola, T. K., Luk, D. & Curran, T. *Molec. cell. Biol.* **13**, 3782-3791 (1993).
26. Bengali, E. *et al. Cell* **68**, 507-519 (1992).
27. Li, L., Chambard, J.-C., Karin, M. & Olson E. N. *Genes Dev.* **6**, 676-689 (1992).
28. Valge-Archer, V. E., de Villiers, J., Sinskey, A. J. & Rao, A. *J. Immun.* **145**, 4355-4364 (1990).

ACKNOWLEDGEMENTS. We thank T. Badalian for technical support, and T. Soderling and B. Perrino for calcineurin inhibitory peptide. This work was supported by the Medical Foundation (J.J.), the Leukemia Society of America (P.G.M.), the American Cancer Society (Z.M.), the Heien Hay Whitney Foundation (T.K.K.), the Institute of Chemistry and Medicine funded by Hoffmann-La Roche Inc. (J.N.L., G.L.V.) and the NIH (A.R.).

An alternative pathway for transcription initiation involving TFII-I

Ananda L. Roy, Sohail Malik*, Michael Meisterernst* & Robert G. Roeder

Laboratory of Biochemistry and Molecular Biology,
The Rockefeller University, 1230 York Avenue, New York,
New York 10021-6399, USA

THE minimal promoter elements required for initiation by RNA polymerase II include the TATA box and/or an initiator element (Inr) at or near the transcription start site¹⁻⁴. Studies of the adenovirus major late core promoter (containing both elements) have demonstrated an initiation pathway that involves binding of the transcription factor TFIID (or the derived subunit, the TATA-binding protein TBP (TFIID_T)) to the TATA element, which is facilitated by transcription factor TFIIA, followed by sequential interactions of other general factors⁵⁻⁸. Here we describe a novel pathway that requires an intact Inr and the Inr-binding factor TFII-I (ref. 3). Sequential addition of the general factors generated TFII-I-dependent preinitiation complexes different from those formed with TFIIA. Furthermore, TBP bound cooperatively (with only TFII-I) to an Inr-containing TATA-less promoter, suggesting a means for activation of TATA-less promoters, which nonetheless require TFIID (refs 9-11). These observations provide support for functionally distinct pathways which could be subject to differential regulation by specific activators or repressors.

TFII-I can bind independently to Inr elements and stimulate basal transcription from the adenovirus major late (AdML) core promoter³. Figure 1a shows further that although either TFII-I or TFIIA support transcription from an intact core promoter containing both TATA and Inr elements (lanes 1-4), only TFIIA could support transcription from an Inr mutant template previously shown³ to be incapable of binding TFII-I (lanes 5-8).

Received 25 February; accepted 29 June 1993.

1. Schreiber, S. L. & Crabtree, G. R. *Immun. Today* **13**, 136-142 (1992).
2. Walsh, C. T., Zydowsky, L. D. & McKeon, F. D. *J. biol. Chem.* **267**, 13115-13118 (1992).

* Present addresses: Medical Research Division, Lederle Laboratories, Pearl River, New York 10965, USA (S.M.); Laboratory for Molecular Biology, Max Planck Institute, University of Munich, 8033 Martinsried, Germany (M.M.).

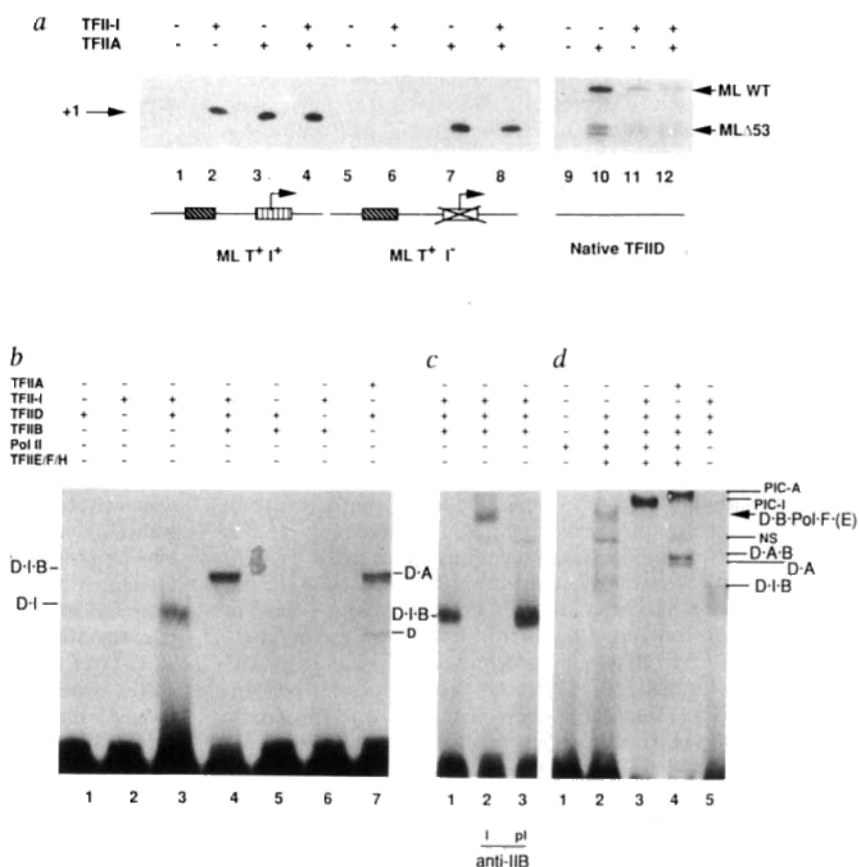
These results indicate that TFII-I function requires direct interactions with the Inr. Furthermore, simultaneous addition of TFII-I and TFIIA (lane 4) did not alter the transcription efficiency, suggesting that these factors act through and compete for a common limiting component of the general transcription machinery. As TBP normally functions on class II promoters as part of the multisubunit TFIID, these analyses were repeated with native TFIID in place of TBP. But even under these conditions there was an additional factor requirement that could be met by either TFIIA or TFII-I (Fig. 1a, lanes 9–12). These results underscore the general importance of TFII-I as a basal transcription factor.

Specific interactions of TFII-I and TBP were studied by electrophoretic mobility shift assay with a probe containing AdML TATA and Inr sequences (Fig. 1b). Under our assay conditions, neither TFII-I (lane 2) nor TBP (lane 1) showed binding, but together (lane 3) they generated a promoter complex (D·I), indicative of highly cooperative interactions between TFII-I and TBP. The mobility of this D·I complex is different from that of a D·A promoter complex¹² (lane 7) and oligonucleotide competition assay verified its specificity (data not shown). Addition of TFIIIB gave rise to a D·I·B complex (lane 4). This complex (Fig. 1c, lane 1) must contain TFIIIB because it was supershifted with a TFIIIB-specific antibody (lane 2) but not with preimmune serum (lane 3). Subsequent addition to the D·I·B complex (Fig. 1d,

lane 5) of RNA polymerase II (pol II) and a TFIIIE/F/H fraction gave rise to a stable preinitiation complex (PIC-I) (lane 3). This complex appears to be structurally distinct from the previously described^{6,7} TFIIA-promoted preinitiation complex (PIC-A) formed with TBP, TFIIIB, pol II, and TFIIIE/F/H (compare lanes 4 and 3). These last components generated a weak promoter complex in the absence of TFIIA or TFII-I (lane 2, minor band indicated by arrow), which could reflect the formation of a previously reported¹³ D·B·Pol·F·(E) complex; but this is probably not functionally significant under our assay conditions as demonstrable transcription requires, in addition to the components leading to its formation, either TFIIA or TFII-I (Fig. 1a).

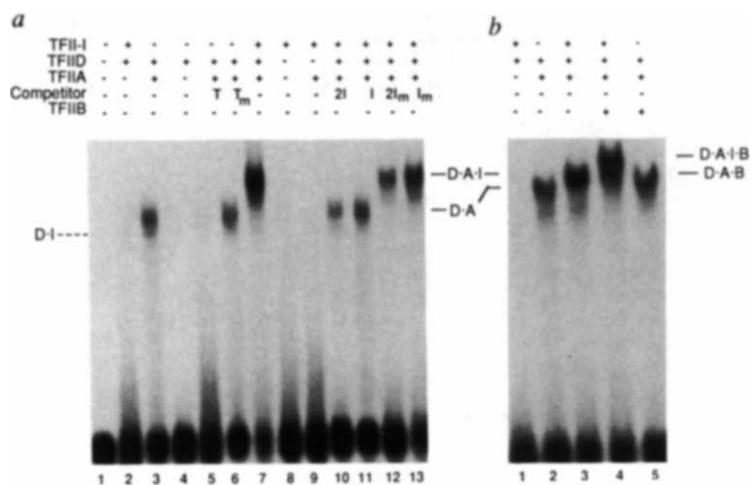
Surprisingly, simultaneous addition of TFII-I and TFIIA to promoter-bound TBP results in a unique but specific D·A·I promoter complex (Fig. 2a, lane 7; see also lanes 10–13 and figure legend) distinct from both the specific D·A promoter complex (lane 3; see also lanes 5 and 6) and the D·I promoter complex (lane 2; formation of the D·I complex was restricted to see more clearly the cooperative nature of D·A·I complex formation). Addition of TFIIIB generated a new D·A·I·B promoter complex (Fig. 2b, lane 4) which was distinct from both the D·A·I promoter complex (lane 3) and the D·A·B promoter complex (lane 5). Thus there is the potential, given adequate concentrations of both TFII-I and TFIIA, for assembly of a complex containing

FIG. 1 TFII-I is an initiator element-binding transcription factor which promotes a novel pathway for preinitiation complex assembly. **a**, TFII-I is a basal initiation factor dependent upon Inr binding for function. Reactions contained common initiation factors TFIIIB, TBP (lanes 1–8) or native partially purified TFIID¹⁵ (lanes 9–12), TFIIIE/F/H, RNA polymerase II (all lanes); AdML wild type (ML T⁺I⁺; lanes 1–4) or Inr mutant (ML T⁺I⁻; lanes 5–8) templates; and TFII-I (lanes 2, 6, 11) or TFIIA (lanes 3, 7, 10) or both (lanes 4, 8, 12). The wild-type template contained AdML promoter sequences from -45 to +65, including both Inr 1 (-3 to +9) and Inr 2 (+45 to +57), fused to the chloramphenicol acetyltransferase CAT gene. The Inr mutant template contained the same sequences apart from mutations in both Inr elements. The normal (+1) transcription start site is indicated by an arrow. The G-less cassettes have been described¹⁵. ML WT, AdML wild type; MLΔ53, AdML sequences from -53 to +10. **b**, Cooperative interactions between TFII-I and TBP, and preinitiation complex assembly on a wild-type AdML core promoter containing both TATA and Inr elements. DNA-protein interactions were analysed by electrophoretic mobility shift assay (EMSA) with a probe containing AdML promoter sequences from -50 to +45. Reactions containing recombinant or partially purified factors, TBP and TFIIIB (lane 5) or TFII-I and TFIIIB (lane 6) failed to yield any complexes, but TFIIA and TBP gave a D·A complex (lane 7). **c**, TFIIIB in the D·I·B promoter complex. Reactions containing TBP, TFII-I and TFIIIB were incubated with no additions (lane 1), with a TFIIIB-specific immune serum (I, lane 2) or with preimmune serum (PI, lane 3). **d**, Formation of TFII-I-dependent (PIC-I) and TFIIA-dependent (PIC-A) preinitiation complexes with other general factors and RNA polymerase II. Factors were added as indicated. A reaction containing the D·I·B promoter complex (lane 5) is shown as a reference lane. The complexes designated PIC-I and PIC-A (lanes 3 and 4) needed all of the factors indicated. The band designated NS (and other unmarked bands) resulted from nonspecific interactions of the RNA polymerase II preparation with the probe. **METHODS**. Reactions contained, in order of addition: 50 ng wild-type or mutant template; TFII-I (10 ng) or TFIIA (100 ng); TBP, TFIID, TFIIIB,



TFIIIE/F/H and RNA polymerase II as given previously¹⁵. The mutant template contained alterations (+3 T→A, +4 C→A, +5 T→A in Inr 1, and +42 G→C, +46 G→C, +48 A→G, +53 C→G, +57 T→A and +58 G→C in Inr 2) known to inhibit TFII-I binding³. The transcribed RNA was scored by primer extension analysis as described¹⁷. EMSA reactions contained 20 fmol wild-type AdML probe (-50 to +45), a highly purified fraction of TFII-I (10 ng) or TFIIA (100 ng), recombinant TBP (0.5 ng), recombinant TFIIIB (5 ng) and partially purified RNA pol II and TFIIIE/F/H in buffer C^{7,12}.

FIG. 2 Formation of promoter complexes containing TFII-I, TFIIA and TBP on a wild-type AdML promoter. *a*, Formation of a D-A-I complex on the AdML core promoter (−50 to +45). Complex formation was monitored by EMSA, with factors and competitor oligonucleotides added as indicated. Oligonucleotides were added in either 40-fold (lanes 10, 12) or 20-fold (lanes 5, 6, 11, 13) molar excess; 3 ng TFII-I was used. Although the D-I promoter complex was visible in this analysis (lane 2), its formation was restricted (by using less TFII-I) in order to visualize the formation of a specific D-A-I promoter complex better (as this otherwise would obscure the upper part of the gel). Involvement of the TATA element in formation of the D-A complex was indicated by sensitivity to wild-type (lane 5) but not mutant (lane 6) TATA-containing oligonucleotides. The specificity of the D-A-I complex, and a requirement for Inr- and TATA-binding proteins, was demonstrated by sensitivity to intact, but not mutant, Inr-containing (lane 10–13) and TATA-containing (data not shown) unlabelled oligonucleotides. Oligonucleotide competitors: wild-type TATA (T), containing AdML sequences (multimerized) from −40 to +15; mutant TATA (Tm), equivalent to wild type apart from mutations changing TATA to GAGA; Inr (I), containing AdML sequences from −15 to +23; Inr mutant (Im), equivalent to wild type except for changes (+3 T→A, +4 C→A, +5 T→A) in Inr 1. *b*, Formation of a putative D-A-I-B promoter complex. Complex formation was monitored by EMSA, with factors added as indicated. Under limiting



concentrations of TFII-I, both D-I complex formation (lane 1) and D-I-B complex formation (data not shown) were restricted. Direct comparative studies have shown that a D-I-B complex migrates slightly faster than the D-A-B complex (data not shown).

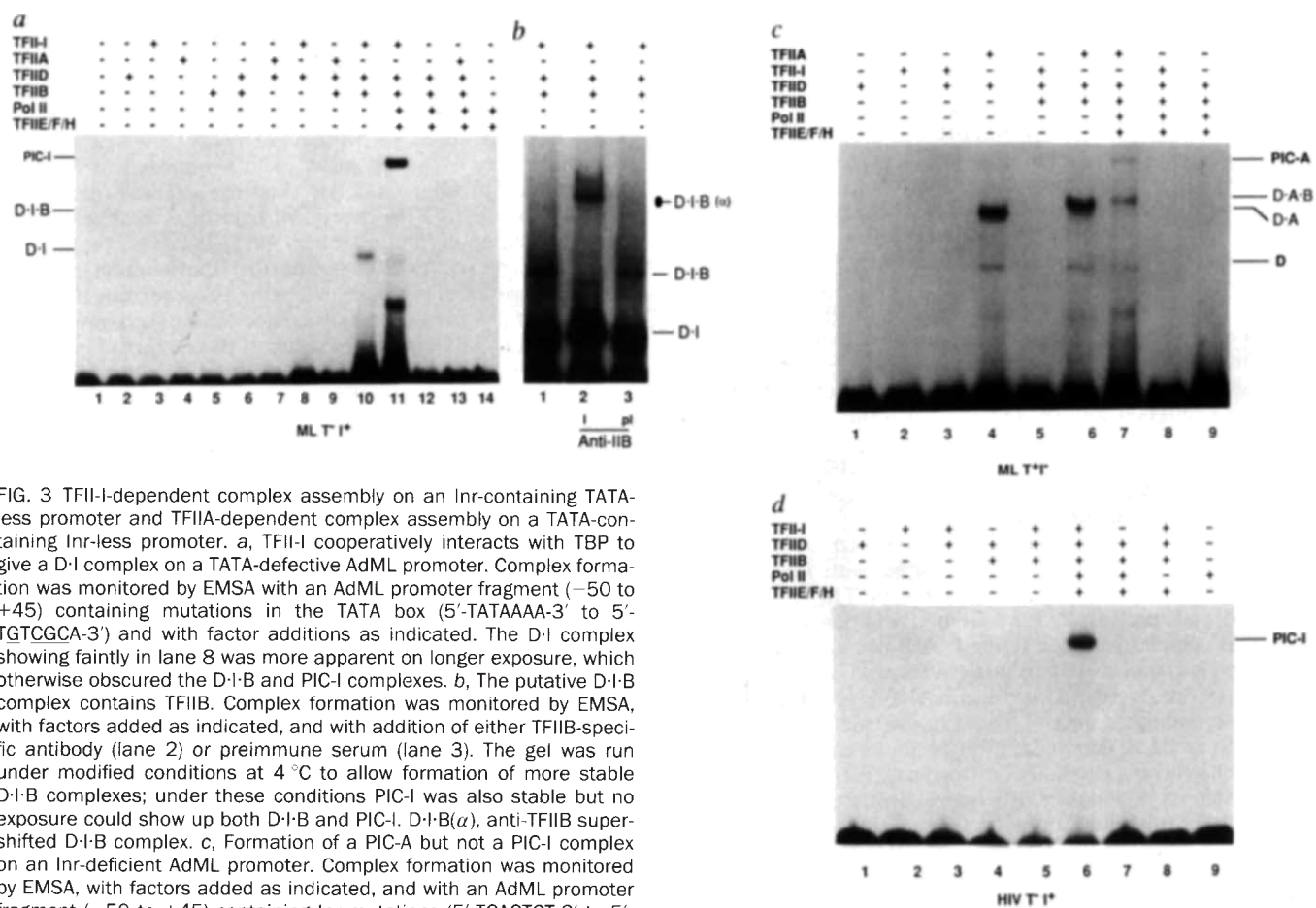
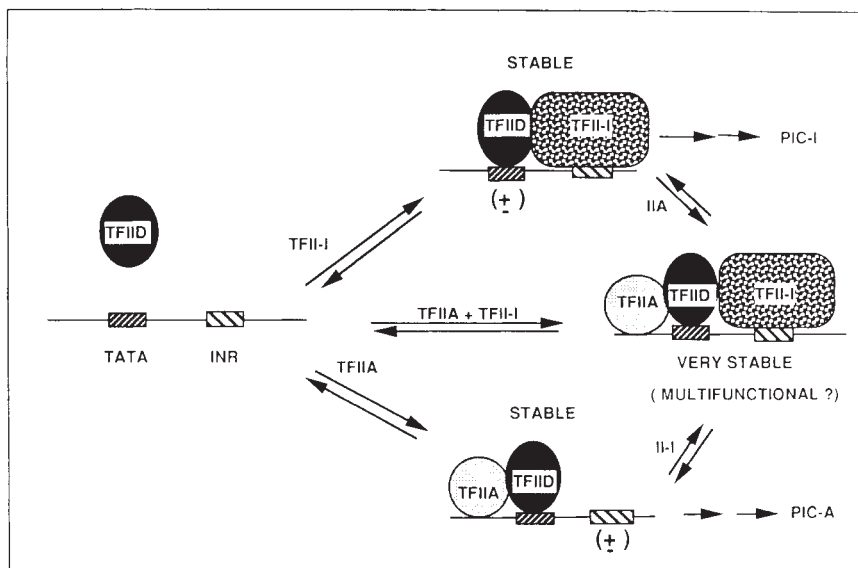


FIG. 3 TFII-I-dependent complex assembly on an Inr-containing TATA-less promoter and TFIIA-dependent complex assembly on a TATA-containing Inr-less promoter. *a*, TFII-I cooperatively interacts with TBP to give a D-I complex on a TATA-defective AdML promoter. Complex formation was monitored by EMSA with an AdML promoter fragment (−50 to +45) containing mutations in the TATA box (5′-TATAAAA-3′ to 5′-TGTCGCA-3′) and with factor additions as indicated. The D-I complex showing faintly in lane 8 was more apparent on longer exposure, which otherwise obscured the D-I-B and PIC-I complexes. *b*, The putative D-I-B complex contains TFII-B. Complex formation was monitored by EMSA, with factors added as indicated, and with addition of either TFII-B-specific antibody (lane 2) or preimmune serum (lane 3). The gel was run under modified conditions at 4 °C to allow formation of more stable D-I-B complexes; under these conditions PIC-I was also stable but no exposure could show up both D-I-B and PIC-I. D-I-B(α), anti-TFII-B super-shifted D-I-B complex. *c*, Formation of a PIC-A but not a PIC-I complex on an Inr-deficient AdML promoter. Complex formation was monitored by EMSA, with factors added as indicated, and with an AdML promoter fragment (−50 to +45) containing Inr mutations (5′-TCACTCT-3′ to 5′-TCACGGG-3′) that restrict TFII-I binding³. *d*, Formation of PIC-I on a TATA-less HIV-1 promoter²². Complex formation was monitored by EMSA. The mutant TATA HIV fragment used as probe was generated by cutting the plasmid pH-GC²² by *Bam*HI and *Bgl*III (−167 to +33) and labelling with

Klenow as described³. Nonspecific binding of pol II was not seen with this fragment (lane 9), although it was with an oligonucleotide probe (Fig. 1d, lane 2). No TFIIA-dependent PIC-A complex (compare *c* with Fig. 1d) was formed under the same conditions (data not shown).

FIG. 4 Model for alternative preinitiation complex (PIC) assembly pathways involving distinct core promoter initiation factors. On composite core promoters containing both TATA and Inr elements, TFIID interactions at the TATA element may be stabilized by sequence-independent interactions with TFIIA, or by Inr-dependent interactions with TFII-I, or by interactions with both TFIIA and TFII-I. These interactions could lead to pathways and complexes that respond selectively to different regulators, or to a composite pathway with a multifunctional complex responsive to a broader class of regulators. Complex formation by the TFIIA-dependent pathway would be favoured on TATA-containing promoters lacking Inr elements, whereas complex formation by the TFII-I-dependent pathway would be favoured on TATA-less promoters containing Inr elements.



both factors. Should TFIIA and TFII-I mediate differential responses to regulatory factors, the latter pathway might show a broader response by virtue of a multifunctional complex.

In view of the cooperativity between TFII-I and TBP, we investigated a possible role of TFII-I in recruiting TBP to Inr-containing promoters lacking a TATA element (Fig. 3a). With a mutant AdML core promoter lacking the TATA element, and under conditions in which neither TBP (lane 2) nor TFII-I (lane 3) alone showed stable binding, both factors together resulted in a specific D·I promoter complex (lane 8; see figure legend). Under the same conditions, and consistent with previous results^{6,7,12}, no comparable complex was formed either with TFIIA alone (lane 4) or with TFIIA and TBP (lane 7). But addition of TFIIIB with TBP and TFII-I gave rise to D·I·B (as well as D·I) promoter complexes (lane 10), whereas no complexes were observed upon addition of TFIIIB either alone (lane 5) or with TFIIA (lane 6), or with TBP plus TFIIA (lane 9). The presence of TFIIIB in the D·I·B complex (Fig. 3b, lane 1) was demonstrated by supershift with a TFIIIB-specific antibody (lane 2) but not with preimmune serum (lane 3). Although the extent of D·I·B complex formation was low relative to that with the intact (TATA⁺Inr⁺) promoter, indicating an intrinsic instability of the D·I·B complex in the absence of TATA, a stable preinitiation complex (PIC-I) was formed upon addition of other general factors (Fig. 3a, lane 11). Formation of this complex was dependent upon addition of TFII-I (lane 12). No similar TFIIA-dependent stable preinitiation complex was seen under the same conditions (lane 13).

In contrast, on a TATA-containing promoter lacking an Inr (Fig. 3c), TFII-I did not form any complex with TBP (lane 3), whereas a TFIIA-dependent complex with TBP was formed (lane 4) as expected^{13,14}. Furthermore, addition of TFIIIB to this D·A complex gave rise to the D·A·B complex (lane 6); no comparable complex (D·I·B) was evident with TFII-I and TBP (lane 5). Consistent with our functional analysis (Fig. 1a) and previously published data^{7,13}, only PIC-A (lane 7) and no PIC-I (lane 8) could be detected on this promoter upon addition of the other general factors. Taken together, these data indicate that formation of a stable TFII-I-dependent preinitiation complex with TBP requires an intact Inr, and that formation of a TFIIA-promoted complex does not. Furthermore, the cooperative interactions between TBP, TFII-I and TFIIA on other promoters are insufficient to form stable D·I·A complexes on promoters lacking either TATA or Inr elements (data not shown).

The CTCA motif within the Inr (CTCACTCTCTCC, initiating nucleotide underlined) of the AdML promoter has a weak

potential to form a functional D·B·Pol·F·(E) complex in the absence of the TATA element¹³. To rule out any contribution of this element to TFII-I-dependent complex formation, we analysed an Inr-containing TATA-less promoter (human immunodeficiency virus, HIV, mutated in TATA) without this motif. Although TFII-I can recognize the HIV-1 initiator (CTGGGTCTCTCT, initiating nucleotide underlined) under some conditions³, stable complexes on the HIV TATA⁺Inr⁺ promoter fragment were not observed under our conditions with any combination of TBP, TFII-I and TFIIIB (Fig. 3d, lanes 1–5). But addition of TFIIIE/F/H and pol II with these factors generated a stable promoter complex (lane 6) which was completely dependent upon TFII-I (lane 7), pol II (lane 8) and TBP (data not shown). The specificity of this complex was further demonstrated by oligonucleotide challenge analysis (data not shown). Thus the TFII-I-dependent complexes represent primary recognition of the Inr element by TFII-I.

In summary, our results indicate that transcription can be initiated by either of two assembly pathways generating distinct preinitiation complexes, or by a composite pathway involving stabilization of TBP binding by either of two factors (TFIIA or TFII-I) or by both factors (Fig. 4). The actual pathway for a given promoter may depend partly on the resident core promoter elements, partly on the relative concentrations of TFIID, TFIIA and TFII-I, and partly on the presence of regulatory factors. Alternative pathways may provide a means of varying the response to specific regulatory factors and hence another level of transcriptional control. For example, the constraints on promoter function imposed by certain negative cofactors^{15,16} may not be possible in the TFII-I-dependent pathway because these negative factors interact with TBP in competition with TFIIA^{12,15}; such cofactors could either switch transcription initiation on a given promoter from a TFIIA-dependent pathway to a TFII-I-dependent pathway or be ineffective on a class (such as TATA-less) of promoters operating mainly by the TFII-I-dependent pathway. TFII-I-dependent initiation also may be regulated by interaction with the helix-loop-helix proteins USF^{3,17} and c-Myc¹⁸. Likewise, there are indications of distinct initiator elements and corresponding interacting factors in other promoters¹⁹. □

Received 29 March; accepted 16 August 1993.

1. Grosschedl, R. & Birnstiel, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1432–1436 (1980).
2. Smale, S. T. & Baltimore, D. *Cell* **57**, 103–113 (1989).
3. Roy, A. L., Meisterernst, M., Pognonec, P. & Roeder, R. G. *Nature* **354**, 245–248 (1991).

4. Roeder, R. G. *Trends biochem. Sci.* **16**, 402–408 (1991).
5. Van Dyke, M., Roeder, R. G. & Sawadogo, M. *Science* **241**, 1335–1338 (1988).
6. Buratowski, S., Hahn, S., Guarente, L. & Sharp, P. A. *Cell* **56**, 549–561 (1989).
7. Maldonado, E., Ha, I., Cortes, P., Weis, L. & Reinberg, D. *Molec. cell. Biol.* **10**, 6335–6347 (1990).
8. Zewel, L. & Reinberg, D. *Current. Opin. Cell Biol.* **4**, 488–495 (1992).
9. Smale, S. T., Schmidt, M. C., Berk, A. J. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4509–4513 (1990).
10. Pugh, B. F. & Tijian, R. *Genes Dev.* **5**, 1935–1945 (1991).
11. Sharp, P. A. *Cell* **68**, 819–821 (1992).
12. Meiternst, M. & Roeder, R. G. *Cell* **67**, 557–567 (1991).
13. Carcamo, J., Buckbinder, L. & Reinberg, D. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9999–10003 (1991).
14. Killen, M., Coulombs, B. & Greenblatt, J. J. *biol. Chem.* **267**, 9463–9466 (1992).
15. Meisterernst, M., Roy, A. L., Lieu, H. L. & Roeder, R. G. *Cell* **66**, 981–993 (1991).
16. Inostroza, J. A. et al. *Cell* **70**, 477–490 (1992).
17. Du, H., Roy, A. L. & Roeder, R. G. *EMBO J.* **12**, 501–511 (1993).
18. Roy, A. L., Carruthers, C., Gutjahr, T. & Roeder, R. G. *Nature* **365**, 359–361 (1993).
19. Seto, E., Shi, Y. & Shenk, T. *Nature* **354**, 241–245 (1991).
20. Olsen, H. S. & Rosen, C. J. *Vir.* **66**, 5594–5597 (1992).

ACKNOWLEDGEMENTS. We thank T. Gutjahr, E. Moorefield and C. Novina for critically reading the manuscript, D. K. Lee for AdML oligonucleotide probes, H. Du for the AdML-CAT construct, C. Rosen for the HIV plasmid (pH-GC), C. Parada and members of our laboratory for critical comments and support. A.L.R. is a Damon Runyon-Walter Winchell Cancer Fund postdoctoral fellow. This study was supported by NIH grants to R.G.R. and by general support from the Pew Trusts to The Rockefeller University.

Direct role for Myc in transcription initiation mediated by interactions with TFII-I

Ananda L. Roy, Cynthia Carruthers, Thomas Gutjahr & Robert G. Roeder

Laboratory of Biochemistry and Molecular Biology,
The Rockefeller University, 1230 York Avenue, New York,
New York 10021, USA

THE nuclear proto-oncoprotein Myc has been implicated in the control of cell proliferation and differentiation^{1–3}. Myc participates in transcription^{4,5} and belongs to the basic-helix-loop-helix (bHLH) family of regulatory proteins^{2,3}. Here we show that Myc interacts with TFII-I, a transcription initiation factor that activates core promoters through an initiator element (Inr)⁶. As previously observed for the bHLH activator USF⁶, Myc was found to interact cooperatively with TFII-I at both Inr and upstream E-box promoter elements. However, in this case Myc interactions with TFII-I at the Inr lead to an inhibition of transcription initiation. This inhibition is selective for a TFII-I-dependent (as opposed to TFIIA-dependent) initiation pathway and correlates with the prevention of complex formation between the TATA-binding protein TBP (TFIID₇), TFII-I and the promoter. TBP probably interacts with Myc, but only slowly. These observations indicate that Myc has the potential to interact physically and functionally with components of the general transcription machinery.

Because Myc and USF both belong to the bHLH-zip family of transcription factors^{7,13} and recognize the same E-box sequences (CACGTG)^{7,8,14}, we investigated whether Myc and TFII-I could interact physically and functionally. Physical interactions at Inr and E-box sites on the adenovirus major late (AdML) promoter were analysed by electrophoretic mobility shift assay. Using an E-box probe (Fig. 1a), full-length Myc alone did not bind^{14–16} under our conditions (lane 4), whereas TFII-I generated specific complexes⁶ (lane 2; see also figure legend). TFII-I and Myc together generated a new, specific complex (lane 3). The presence of Myc was indicated by a specific supershift of the complex with anti-Myc antibody (lane 8, arrow) but not pre-immune serum (lane 9).

A similar analysis with an Inr-containing probe (Fig. 1b) revealed binding of TFII-I alone (lane 2); full-length Myc alone (lane 1) did not bind. But TFII-I and Myc together (lane 3)

again generated a new, specific complex (marked with a circle), which presumably arises from simultaneous binding of TFII-I and Myc to the promoter. The specificity of the complexes was also demonstrated by competition with intact (lane 4) but not mutant (lane 5) Inr oligonucleotides. Interactions between TFII-I and Myc were also detected by co-immunoprecipitation with a monoclonal antibody against Myc following incubation of recombinant Myc and partially purified TFII-I (Fig. 1c; also see figure legend). The specificity of these interactions is indicated by their dependence on C-terminal Myc domains, which are implicated in other protein-protein interactions; thus they are eliminated by the Δ H/LZ and Δ HLH mutations but not by an extensive N-terminal deletion (C176) in Myc (Fig. 1c). Furthermore, a little TFII-I co-precipitated with Myc even in the presence of a stoichiometric excess of Max (data not shown). These interactions could therefore be important under appropriate physiological conditions.

The Myc and TFII-I interaction at the AdML Inr element provided a rationale for analysing any effect of Myc upon core promoter function that is dependent upon TFII-I. TFII-I and TFIIA independently promote transcription initiation by RNA polymerase II in the presence of the other common factors^{17,18} and so provide two alternative pathways for *in vitro* activation of the AdML core promoter. When analysed in the respective reconstituted systems¹⁹ (Fig. 2), Myc specifically inhibits TFII-I-dependent transcription in a dose-dependent manner (lanes 1–5), but does not affect TFIIA-dependent transcription even at high concentration (lanes 8 and 9). Consistent with an inhibition involving stable Myc-TFII-I interaction, increasing TFII-I relieves the inhibitory effect at a fixed Myc concentration (lanes 5–7). That this interaction is specific is demonstrated by the failure of Δ H/LZ and Δ HLH deletion mutants of Myc to inhibit transcription, even at levels fourfold higher than those (50 ng) at which wild-type Myc completely abolishes TFII-I-dependent transcription (Fig. 2b). Furthermore, these effects are mediated through the core promoter elements, as indicated (Fig. 2b) by inhibition of both wild-type and Δ 53 (minus the upstream USF binding site) AdML templates²⁰. These results indicate that interactions between Myc and TFII-I are specific and functional and further distinguish alternative pathways for initiation: a TFII-I-dependent pathway inhibited by Myc versus a TFIIA-dependent pathway relatively refractory to Myc.

Figure 3a and b shows that Myc inhibits the formation of TBP-TFII-I promoter complexes (compare lanes 2 and 1 in Fig. 3a and b) and that this effect is relieved by either an excess of TFII-I (Fig. 3a, lanes 3 and 4) or a Myc-specific antiserum (Fig. 3b, lanes 3 and 4). In contrast, Myc does not affect the formation of TFIIA-TBP promoter complexes (Fig. 3b, lanes 5–7). This suggests that Myc inhibits TFII-I-dependent transcription by interfering with the formation of TFII-I-TBP promoter complexes, whereas TFIIA-dependent transcription is refractory to Myc as a result of the insensitivity (to Myc) of TFIIA-TBP promoter complexes. But preincubations of Myc and TBP (Fig. 3c) before adding probe and either TFII-I (lane 2 versus lane 1) or TFIIA (lane 6 versus lane 5) prevented formation of TBP-TFIIA or TBP-TFII-I promoter complexes. As before, these inhibitory effects of Myc are reversed by a Myc-specific antiserum (lanes 3, 4 and 7). Myc may therefore interact, but more slowly, with TBP as well. Our failure to detect TBP-Myc promoter complexes may be a result of their instability under our electrophoretic conditions or because Myc-TBP interactions only occur off the DNA.

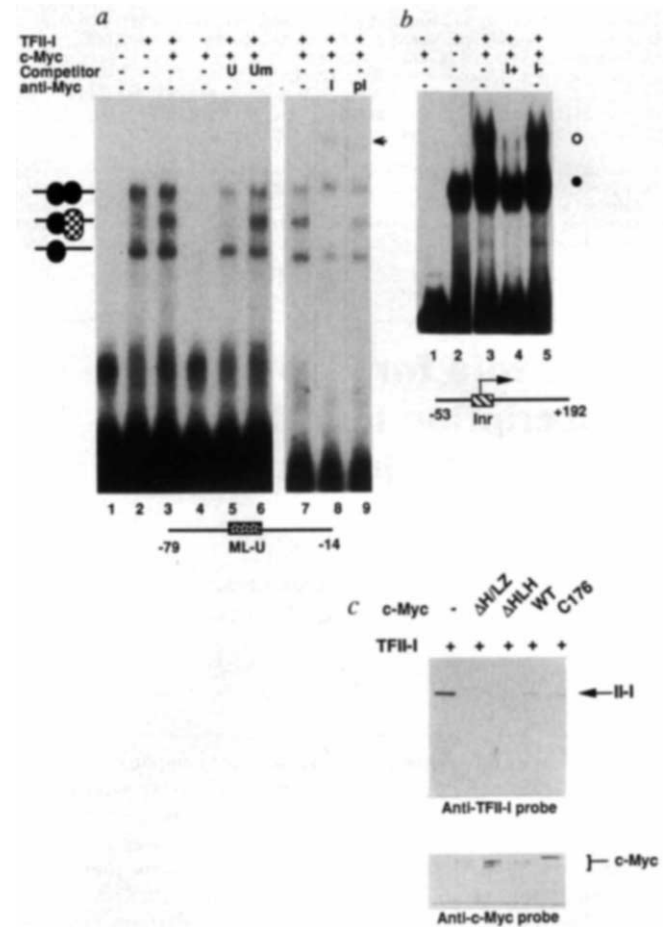
We have shown that Myc can directly affect transcription initiation by its associations with the general transcriptional machinery at the core promoter. We also have identified the Inr-binding factor TFII-I, and possibly TFIID, as potential targets (modelled in Fig. 4). Myc may therefore have other partners or intracellular targets in addition to Max^{15,16,21,22} and Rb²³, although an *in vivo* association between Myc and TFII-I (or TBP) remains to be demonstrated. Overexpressed Myc is known

to activate artificial reporter genes harbouring CACGTG motifs but lacking an Inr element in mammalian⁴ and yeast cells⁵, but our *in vitro* analysis indicates that Myc could have a profound, and opposite, influence on a promoter containing a functional Inr but lacking a cognate E-box binding site. Although the specific inhibitory effects of Myc reported here have not been assayed under native TFIIID-dependent transcription conditions, they are consistent with reports of transcriptional repression by Myc

*in vivo*²⁴⁻²⁶. More significantly, recent transfection studies have shown that Myc can efficiently block TFII-I- and Inr-dependent activation of a natural cyclin D1 promoter that lacks a consensus TATA element (M. Eilers, A.L.R. and R.G.R., unpublished observations). But it is also possible that promoters under the influence of regulatory elements such as upstream E boxes and associated factors might be activated by Myc^{4,5,27} in conjunction either with TFII-I or with another general initiation factor such

FIG. 1 Interactions between TFII-I and Myc at the E-box (ML-U) and initiator (Inr) sites. **a**, TFII-I interacts specifically with Myc at an E-box site. Interactions were monitored by electrophoretic mobility shift assay (EMSA). The ML-U site contains the CACGTG E-box sequence of the AdML promoter which specifically recognizes homomeric or heteromeric Myc complexes¹⁴⁻¹⁶. TFII-I generated two complexes (lanes 2 and 7) previously characterized as C1 and C2 (ref. 6). TFII-I and Myc (lanes 3, 6, 7) generated a new heteromeric complex. There was no interaction of c-Myc in the absence of TFII-I even at high concentrations of Myc (lane 4, and data not shown). Oligonucleotide competitors⁶ at 10-fold molar excess were: wild-type ML-U (U), AdML nucleotides -75 to -35; mutant ML-U (Um), same as wild-type ML-U but with CACGTG changed to TATGCG. **b**, Interactions between TFII-I and Myc at the Inr element. Interactions were monitored by EMSA using a long Inr AdML probe (below) containing Inr 1 and Inr 2 elements⁶. Heterogeneity in the complex seen with TFII-I alone (filled circle) may reflect multiple interactions at the two Inr elements. The novel complex (circle) generated with TFII-I plus Myc is presumed to arise from heteromeric interactions and is seen only at relatively high Myc concentrations, precluding the use of anti-Myc antibodies. Oligonucleotide competitors⁶ added at 10-fold molar excess were: wild-type Inr (I'), AdML nucleotides -15 to +23; mutant Inr (I''), same as wild type, but with mutations T to A at +3, C to T at +4, and T to A at +5. **c**, Physical interactions between Myc and TFII-I. Coimmunoprecipitation of mixtures were monitored with an anti-Myc monoclonal antibody (Oncogene Sciences) followed by western blot analysis using an affinity-purified anti-TFII-I antibody. A control reaction (lane 1) contained only a highly purified⁶ preparation of TFII-I as reference. All immunoprecipitated reactions contained TFII-I and various mutants or wild-type Myc (all Myc proteins were bacterially expressed and purified); Δ H/LZ deletion mutant (Δ 371-433 deleted in the HLH and leucine zipper (LZ) domains) (lane 2), Δ HLH deletion mutant (Δ 371-416 deleted only for HLH) (lane 3); wild type (lane 4); and C176Myc-GST fusion protein (containing the C-terminal 176 amino acids of Myc fused to glutathione S-transferase (GST)) (lane 5). Arrow indicates the position of the 120K TFII-I polypeptide⁶. The same blot was stripped and reprobed with Myc polyclonal antibody.

METHODS. For EMSA analysis (**a** and **b**), a highly purified preparation⁶ of TFII-I (10 ng) and recombinant Myc (20 ng in **a**, 100 ng in **b**) were incubated at 30 °C for 15 min in buffer A containing 20 mM Tris, pH 7.9, 10% (V/V) glycerol, 10 mM EDTA, 1 mM dithiothreitol (DTT), 80 mM KCl and 500 mM urea. After this, and only where indicated, anti-Myc immune serum or preimmune serum was added and the incubation continued for 15 min at 30 °C. Subsequently, 10-20 fmol of labelled (with Klenow enzyme) DNA probe with 10 μ g ml⁻¹ poly(dA-dT) and various oligonucleotide competitors were added, incubated for 20 min at 30 °C, and the reactions analysed on a 4% non-denaturing polyacrylamide gel. For coimmunoprecipitation (**c**), partially purified TFII-I was mixed with 2 μ g of various Myc proteins in buffer A containing 500 mM urea, incubated at 30 °C for 15 min, diluted to 100 μ l with buffer C without



DTT, and incubated with anti-Myc monoclonal antibody for 1 h at 4 °C. Protein A-Sepharose (Oncogene Sciences) was then added and the reactions incubated for an additional 1 h at 4 °C followed by washes with buffer A containing 0.1% NonidetP40 and 0.1% aprotinin. The precipitate was resuspended in SDS loading buffer, run on 15% SDS-PAGE and, after blotting, probed with an anti-TFII-I antibody⁶ (T.G., A.L.R. and R.G.R., unpublished results). The blot was then washed to strip off TFII-I antibody and reprobed with a polyclonal Myc antibody.

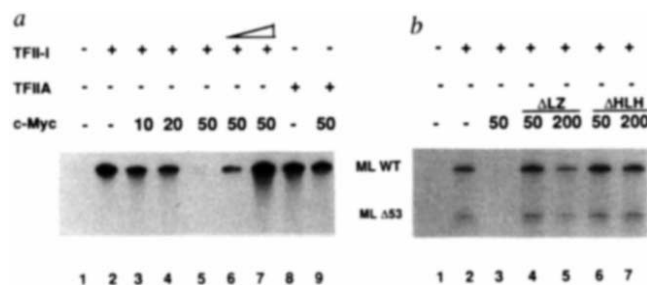
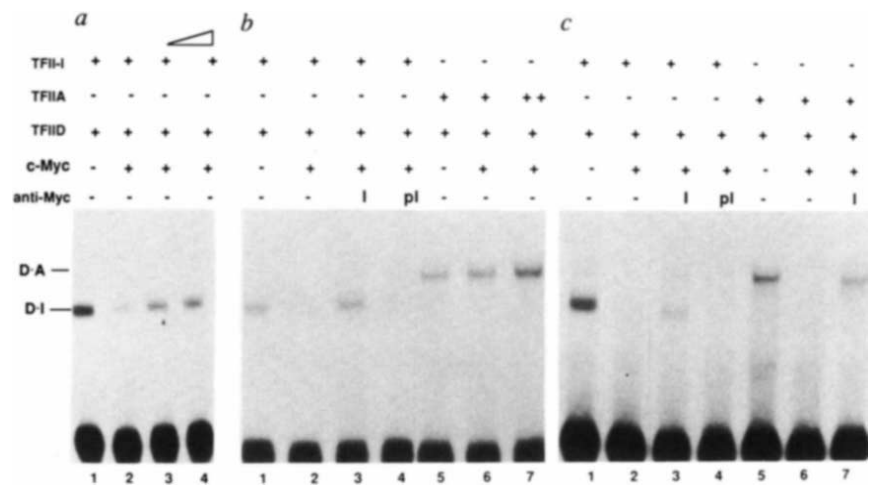


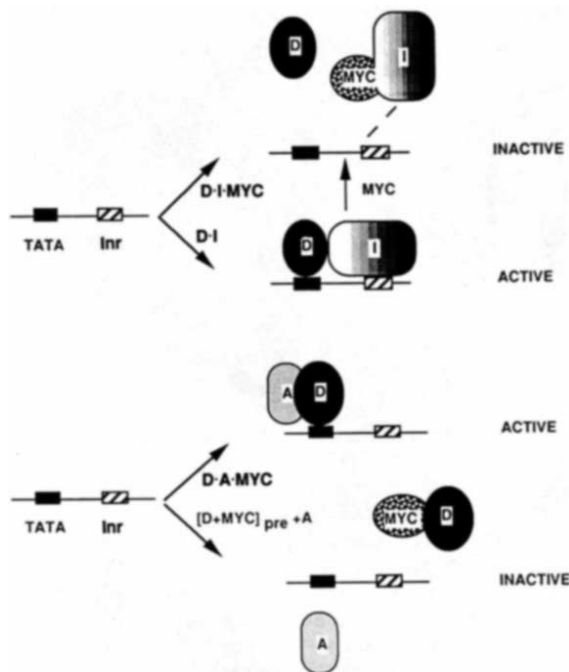
FIG. 2 Myc can directly regulate transcription initiation from a core promoter. All reactions contained common transcription factors TFIIIB, TBP, TFII-E/F/H, RNA polymerase II, either TFIIA or TFII-I (as indicated), the indicated amounts (ng) of Myc, and either **a**, AdML wild-type or **b**, AdML Δ 53 templates²⁰. In **a**, Myc selectively inhibits TFII-I-dependent transcription initiation. Levels of TFII-I in lanes 6 and 7 are 3-fold and 5-fold higher, respectively, than in lanes 2-5. In **b**, Myc specifically inhibits transcription initiation from a core promoter template and requires an intact HLH/LZ domain. Wild-type Myc (lane 3) represses TFII-I-dependent transcription initiation (compare lanes 2 and 3) on both wild-type (WT) and Δ 53 (core promoter only) templates. There was no repression for either comparable (50 ng) or fourfold increased (200 ng) levels of Δ H/LZ deletion mutant (lanes 4, 5) and Δ HLH (lanes 6, 7). **METHODS.** Reaction conditions for *in vitro* transcription were as described^{4,16} except for the inclusion of 25 mM urea. The reactions contained, in order of addition: 50 ng template; TFII-I (10 ng) or TFIIA (100 ng) as indicated; TBP (1 ng); varying amounts (ng) of Myc; TFII-B, TFII-E/F/H and RNA polymerase II in the amounts described¹⁹.

FIG. 3 Myc blocks formation of TBP-TFII-I (D-I) but not TBP-TFII-A (D-A) promoter complexes. *a*, Myc inhibition of D-I complex formation. Interactions were monitored by EMSA with an AdML core promoter probe (−50 to +45) containing a TATA box and an Inr under modified conditions. Under these conditions, TFII-I cooperatively interacts with TBP to generate a stable D-I promoter complex (lanes 1 and 5). Lanes 3 and 4 contained, respectively, 25 ng and 50 ng TFII-I; there is 10 ng TFII-I loaded in both lanes 1 and 2. *b*, D-A promoter complex formation in the presence of Myc. EMSA was done with factors and anti-Myc antibody (I) or pre-immune serum (pl) as indicated. Myc was added at a concentration that inhibited D-I complex formation in *a*. 100 ng (lanes 5, 6) or 250 ng (lane 7) of TFIIA was present. *c*, Preincubation of Myc with TBP prevents formation of both D-I and D-A complexes. EMSA results are shown with TBP preincubated alone (lanes 1, 5) or with TFII-I (lane 1) or TFIIA (lane 5) before addition of other factors and anti-Myc antibody or pre-immune serum.

METHODS. EMSA reaction conditions were as for Fig. 1, except that poly(dG-dC) (2.5 $\mu\text{g ml}^{-1}$) replaced poly(dA-dT) as the nonspecific carrier and reactions contained 5 mM Mg^{2+} . Under these conditions, TFII-I alone fails to give any appreciable promoter complex. The order of addition in *a* and *b* was TBP (1 ng), TFII-I (10 ng) or TFIIA (100 ng), 20 fmol DNA probe and, after incubation on ice for 5 min, Myc (50 ng).



Reactions were then incubated at 30 °C for 20 min and analysed on a 5% polyacrylamide non-denaturing gel containing 25 mM Tris, pH 7.9, 100 mM glycine, 1 mM EDTA, 4 mM Mg^{2+} and 1 mM DTT. For *c*, TBP was preincubated with Myc at 30 °C for 10 min then TFII-I or TFIIA was added, followed by the DNA probe. Antibodies, where indicated, were added with Myc.



as TFIID. Thus the current data lead to the prediction that promoters subject to a natural repression by Myc will contain, minimally, an Inr element that is recognized by TFII-I, whereas promoters that are activated by Myc will lack such elements and/or contain dominant E-box elements. Given our lack of information on the molecular targets and mechanism(s) of action of Myc, the results presented here provide some valuable leads into this problem. □

FIG. 4 Model for potential Myc interactions leading to modulation of transcription initiation. Composite core promoters containing both TATA and Inr elements may be assembled into functional preinitiation complexes either by a TFII-I-dependent pathway that depends upon the Inr, or by a TFIIA-dependent pathway that is independent of an Inr (refs 17, 18). Myc can block formation of a D-I promoter complex, which leads to repression when transcription depends on TFII-I. Stable interactions between TFII-I and Myc probably occur either off or on the Inr element. In either case, TFII-I cannot enter into a productive preinitiation complex. In contrast, Myc and TFIIA interact with TBP in a mutually exclusive fashion, resulting either in a TBP-Myc complex or a TBP-TFIIA complex. Formation of the former leads to repression and the latter to activation. Interaction of the various factors may be determined by their effective concentrations and in part by the kinetics of association (for further discussion, see text). As the stoichiometry of the heteromeric complexes is not known, they are tentatively represented as dimers.

- Pendegast, G. C. & Ziff, E. B. *Trends genet. Sci.* **8**, 91–96 (1992).
- Kretzner, L., Blackwood, E. M. & Eisenman, R. N. *Nature* **359**, 426–429 (1992).
- Amati, B. et al. *Nature* **359**, 423–426 (1992).
- Roy, A. L., Meisterernst, M., Pognonec, P. & Roeder, R. G. *Nature* **354**, 245–248 (1991).
- Gregor, P. D., Sawadogo, M. & Roeder, R. G. *Genes Dev.* **4**, 1730–1740 (1990).
- Pognonec, P. & Roeder, R. G. *Molec. cell. Biol.* **11**, 5125–5136 (1991).
- Murre, C., Schonleber-McCaw, P. & Baltimore, D. *Cell* **56**, 777–783 (1989).
- Murre, C. et al. *Cell* **58**, 537–544 (1989).
- Davis, R. L., Cheng, P. F., Lassar, A. B. & Weintraub, H. *Cell* **60**, 733–746 (1990).
- Jones, N. *Cell* **61**, 9–11 (1990).
- Lamb, P. & McKnight, S. L. *Trends biochem. Sci.* **16**, 417–422 (1991).
- Blackwell, T. K. et al. *Science* **250**, 1149–1152 (1990).
- Blackwood, E. & Eisenman, R. N. *Science* **251**, 1211–1217 (1991).
- Pendegast, G. C., Lawe, D. & Ziff, E. B. *Cell* **65**, 395–407 (1991).
- Roeder, R. G. *Trends biochem. Sci.* **16**, 402–408 (1991).
- Roy, A. L., Malik, S., Meisterernst, M. & Roeder, R. G. *Nature* **365**, 355–359 (1993).
- Meisterernst, M., Roy, A. L., Lieu, H. L. & Roeder, R. G. *Cell* **66**, 981–993 (1991).
- Sawadogo, M. & Roeder, R. G. *Cell* **43**, 165–175 (1985).
- Blackwood, E. M., Luscher, B. & Eisenman, R. N. *Genes Dev.* **6**, 71–80 (1992).
- Wenzel, A. et al. *EMBO J.* **10**, 3703–3712 (1991).
- Rustgi, A. K., Dyson, N. & Bernards, R. *Nature* **352**, 541–544 (1991).
- Van Antwerp, M. E. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 9010–9014 (1992).
- Bernards, R., Dessain, S. K. & Weinberg, R. A. *Cell* **47**, 667–674 (1986).
- Inghirami, G. et al. *Science* **250**, 682–686 (1990).
- Pendegast, G. C. & Cole, M. D. *Molec. cell Biol.* **9**, 124–134 (1989).

ACKNOWLEDGEMENTS. We thank P. Pognonec for bacterially expressed Myc and Myc anti-serum, C. V. Dang for $\Delta\text{H}/\text{LZ}$ and ΔHLH mutants, E. Blackwood and R. Eisenman for C176 Myc protein, B. Moorefield and E. Martinez for critically reading the manuscript, and other members of our laboratory for critical comments and support. A.L.R. is a Damon Runyon-Walter Winchell Cancer Research Fund postdoctoral fellow. T.G. was supported by the Schweizerischer National Fonds. This study was supported by grants from the NIH to R.G.R. and by general support from the Pew Trusts to the Rockefeller University.

Received 29 March; accepted 16 August 1993.

- Cole, M. D. A. *Rev Genet.* **20**, 361–384 (1986).
- Luscher, B. & Eisenman, R. N. *Genes Dev.* **4**, 2025–20435 (1990).

Structure and function of endoglucanase V

Gideon J. Davies*, G. Guy Dodson*, Roderick E. Hubbard*, Shirley P. Tolley*, Zbigniew Dauter†, Keith S. Wilson†, Carsten Hjort‡, Jan Møller Mikkelsen‡, Grethe Rasmussen‡ & Martin Schülein‡

* Department of Chemistry, University of York, Heslington, York YO1 5DD, UK

† EMBL Hamburg Outstation, c/o DESY, Notkestrasse 85, D-2000 Hamburg 52, Germany

‡ Novo Nordisk a/s, Novo Alle, 2880 Bagsvaerd, Denmark

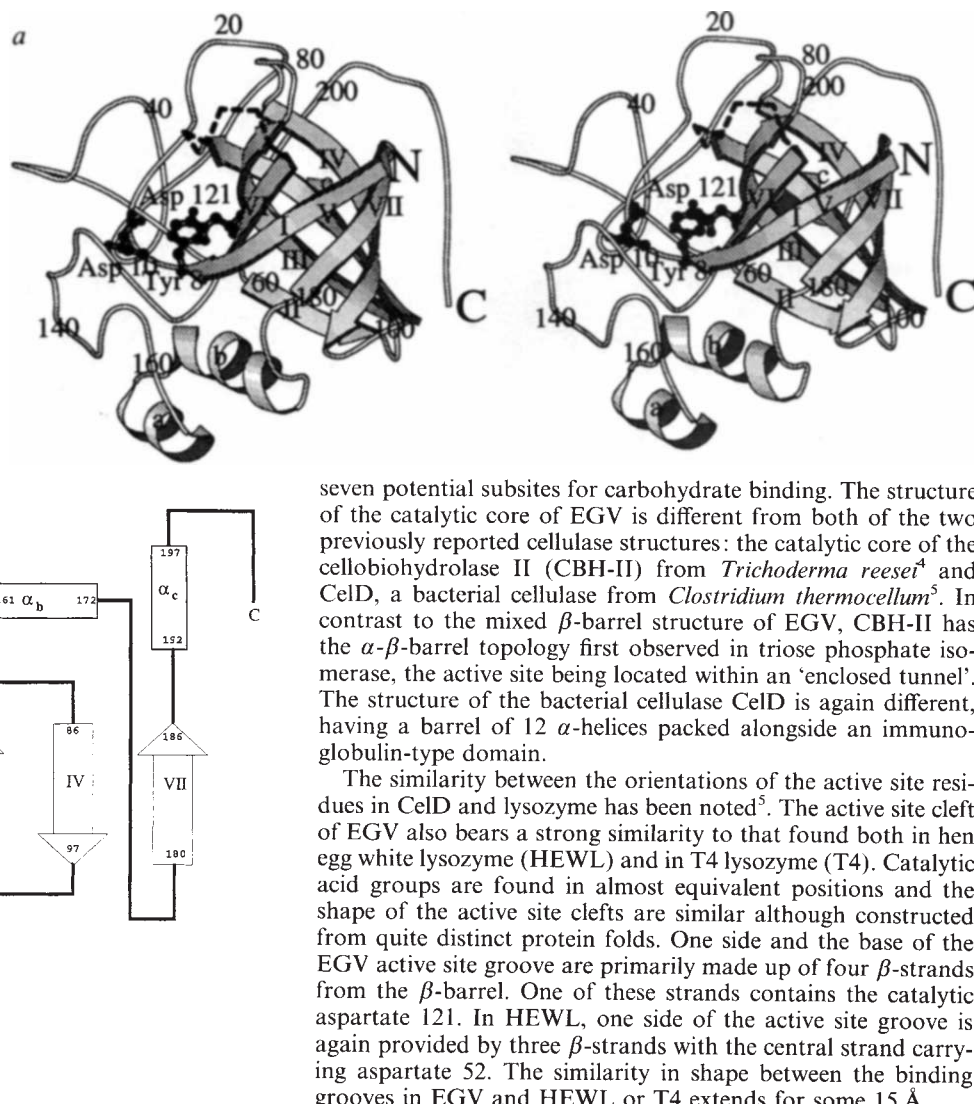
CELLULOSE is the major polysaccharide component of plant cell walls and is the most abundant organic compound on the planet. A number of bacterial¹ and fungal² organisms can use cellulose as a food source, possessing cellulases (cellobiohydrolases and endoglucanases) that can catalyse the hydrolysis of the β -(1,4) glycosidic bonds. They can be classified into seven distinct families³. The three-dimensional structures of members of two of these families are known^{4,5}. Here we report the structure of a third cellulase, endoglucanase V, whose sequence is not represented in any of the above families. The enzyme is structurally distinct from the

previously determined cellulases but is similar to a recently characterized plant defence protein⁶. The active site region resembles that of lysozyme, despite the lack of structural similarity between these two enzymes.

Fungal cellulases typically contain a catalytic core linked to a cellulose-binding domain by a flexible peptide segment. The structure of the catalytic core of an endoglucanase (EGV) from *Humicola insolens* has been determined by X-ray analysis at 1.6 Å resolution. This enzyme is a new endoglucanase whose sequence has no similarity to any in the original Henrissat classification³. The structure has been solved by isomorphous replacement (Table 1). The enzyme has a flattened spheroidal shape with rough dimensions of $42 \times 42 \times 22$ Å. The major structural feature is a six-stranded β -barrel domain. A seventh strand of hydrogen bonds to one of the barrel strands but is not part of the barrel itself. The barrel differs from a standard 'jelly-roll' in that it is composed both of anti-parallel and of parallel β -strands. We believe that this represents a new protein topology (Fig. 1) although the barrel core of the structure is similar to that found in 'barwin', a plant defence protein⁶ (Fig. 2).

A large, deep groove runs across the surface of the molecule, partitioning the β -barrel from the loop regions. Site-directed mutagenesis has identified two aspartates (10 and 121) critical for activity. They are located on either side of the groove with their Ca atoms some 11.5 Å apart. They are positioned above and either side of a tyrosine residue, Tyr 8, which lies at the bottom of the active site groove. The active site groove contains

FIG. 1 a, Stereo ribbon diagram drawn with the MOLSCRIPT¹⁷ program, showing the overall topology of the EGV molecule. The N- and C-terminal residues are indicated. The flexible loop region (residues 112–117), only seen in substrate/inhibitor complexes, is shown dashed. The catalytic aspartic acid residues and the tyrosine, lying at the floor of the binding-site, are shown in 'ball and stick' representation. Secondary structure was determined using the hydrogen-bonding patterns as defined previously¹⁸, as implemented in the QUANTA program (Molecular Simulation Inc., USA). b, Schematic representation illustrating the unusual β -barrel topology of EGV. The flexible loop between barrel strands V and VI is again represented with a dashed line.



seven potential subsites for carbohydrate binding. The structure of the catalytic core of EGV is different from both of the two previously reported cellulase structures: the catalytic core of the cellobiohydrolase II (CBH-II) from *Trichoderma reesei*⁴ and CelD, a bacterial cellulase from *Clostridium thermocellum*⁵. In contrast to the mixed β -barrel structure of EGV, CBH-II has the α - β -barrel topology first observed in triose phosphate isomerase, the active site being located within an 'enclosed tunnel'. The structure of the bacterial cellulase CelD is again different, having a barrel of 12 α -helices packed alongside an immunoglobulin-type domain.

The similarity between the orientations of the active site residues in CelD and lysozyme has been noted⁵. The active site cleft of EGV also bears a strong similarity to that found both in hen egg white lysozyme (HEWL) and in T4 lysozyme (T4). Catalytic acid groups are found in almost equivalent positions and the shape of the active site clefts are similar although constructed from quite distinct protein folds. One side and the base of the EGV active site groove are primarily made up of four β -strands from the β -barrel. One of these strands contains the catalytic aspartate 121. In HEWL, one side of the active site groove is again provided by three β -strands with the central strand carrying aspartate 52. The similarity in shape between the binding grooves in EGV and HEWL or T4 extends for some 15 Å.

TABLE 1 Data collection and phasing statistics

Data collection											
Data	Resolution (Å)	R_{merge}^*		Number of observations		Number of unique measurements		Completeness (%)			
Native	1.6	0.089		102,802		39,421		83			
Cellobiose-complex	1.9	0.071		44,074		12,968		97			
CH ₃ HgCl	3.0	0.053		10,271		5,486		75			
LuCl ₃	2.5	0.045		27,987		10,484		84			
Phasing statistics											
Derivative	$R_{\text{deriv}}^\dagger$ (%)		$R_{\text{cullis}}^\ddagger$		Phasing Power§	Heavy-atom parameters					
	Centric	Acentric	Anomalous			X	Y	Z	B (Å ²)	Occupancy	Anomalous occupancy
CH ₃ HgCl	21.5	0.54	0.66	0.50	1.5	0.025	0.000	0.286	17	0.60	0.54
						0.040	0.033	0.418	17	0.45	0.45
						0.714	0.239	0.568	28	0.30	0.34
LuCl ₃	16.5	0.58	0.65	0.60	1.3	0.089	0.094	0.392	29	0.55	0.33
						0.778	0.249	0.979	28	0.70	0.46

Crystals of native EGv were grown by hanging-drop vapour-phase diffusion with 18–22% PEG-8000 as a precipitant. The initial protein concentration was 12 mg ml⁻¹ in 10 mM Tris-HCl, pH 8.0. Crystals of EGv are in space group $P2_1$ with cell dimensions $a = 41.8$ Å, $b = 51.3$ Å, $c = 44.4$ Å, and $\beta = 109.6^\circ$. There is one molecule of EGv in the asymmetric unit. Derivative screening using the Hendrix–Lentfer image plate at the EMBL Hamburg outstation, with a Mo sealed tube source, identified two sets of derivatives: methyl Hg chloride (2 mM soak, 5 days) and many salts of the Lanthanide series of metals. Native and derivative data were collected at the Photon factory synchrotron facility using the Weissenberg camera. Data were collected at a wavelength of 1.000 Å to reduce absorption errors and optimize the component of the anomalous scattering from the heavy atoms. LuCl₃ (10 mM soak, 10 h) was chosen as, of all the Lanthanide series of metals, it has the greatest values of f'' at low (~1.0 Å) wavelengths. Care was taken with the derivative data to collect Bijvoet pairs on the same image. Screenless Weissenberg data were collected with a camera radius of 286.5 mm, 16° oscillation ranges per image and a coupling constant of 4° mm⁻¹. Data were processed with the WEISS program¹⁰. All other calculations used the CCP4¹¹ suite of programs. Phases were calculated using the MLPHARE¹² program with the phases being further improved using SQUASH¹³. The structure was built using the O program¹⁴ and refined using standard techniques with XPLOR¹⁵ and PROLSQ¹⁶. The crystallographic R -factor is 0.154, at 1.6 Å resolution with r.m.s. deviations from stereochemical target values of 0.010 Å for bonds and 0.025 Å for angles (1–3 bonding distance). Coordinates have been deposited with the Brookhaven Protein Database.

$$* R_{\text{merge}} = \sum_{hkl} |I - \langle I \rangle| / \sum_{hkl} \langle I \rangle.$$

$$\dagger R_{\text{deriv}} = \sum |F_{\text{deriv}} - F_{\text{nat}}| / \sum F_{\text{nat}}.$$

$$\ddagger R_{\text{cullis}} = \sum \left(|F_{\text{PHobs}}| - |F_{\text{PHcalc}}| \right) / \sum \left(|F_{\text{PHobs}}| + |F_{\text{PHcalc}}| \right) \quad \text{for centric or acentric terms, and} \\ \sum \left(|\Delta \text{ano}_{\text{obs}}| - |\Delta \text{ano}_{\text{calc}}| \right) / \sum |\Delta \text{ano}_{\text{obs}}| \quad \text{for anomalous differences.}$$

§ The phasing power is the mean value of the heavy-atom structure amplitude divided by the lack of closure error.

FIG. 2 Stereo figure illustrating the similarity between the β -barrels of 'barwin'⁶ (blue) and EGv (yellow). 38 Ca atoms (representing 18% of the EGv Ca atoms) overlap with the barwin barrel with an r.m.s. deviation of 1.56 Å. Coordinates for barwin (and for T4 and hen-egg white lysozyme) were obtained from the Brookhaven Protein Database¹⁹.

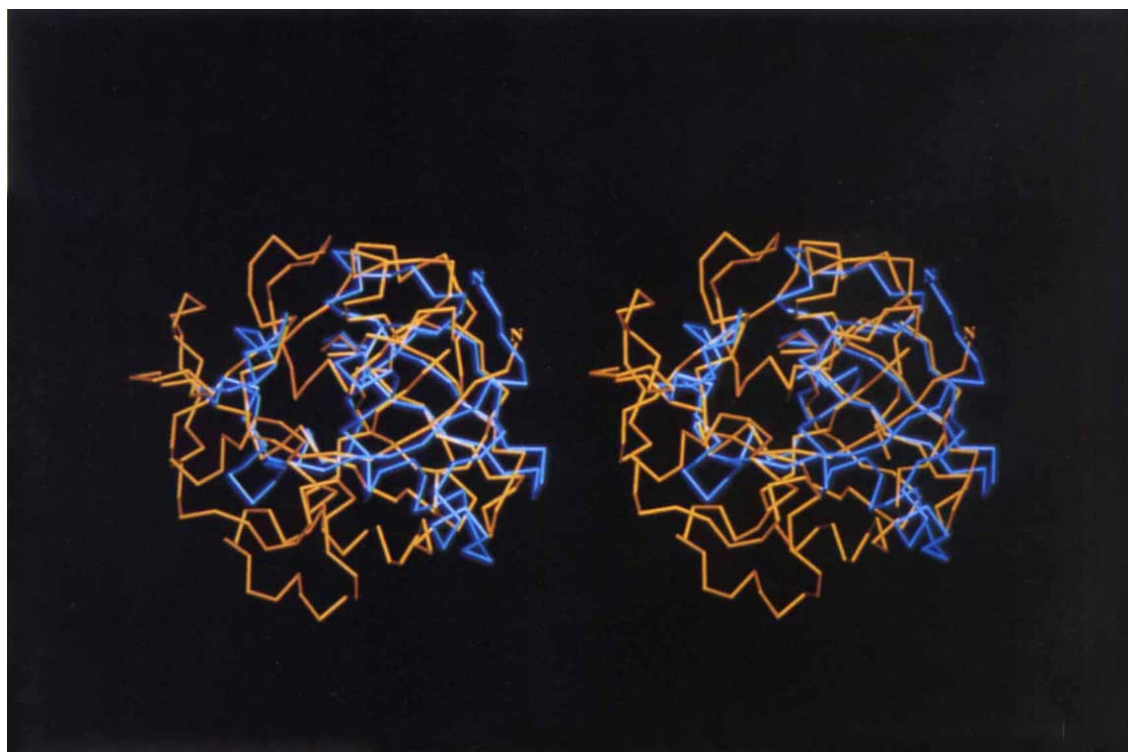
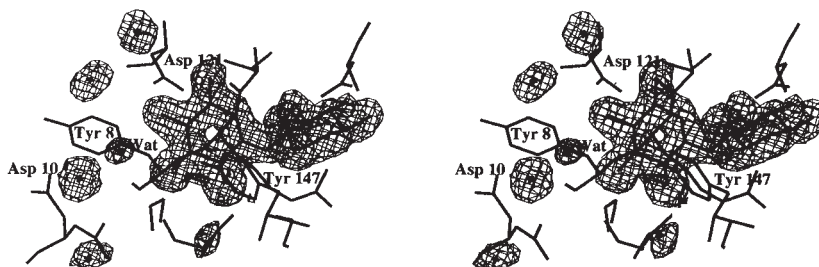


FIG. 3 Stereo figure illustrating the binding of the product cellobiose to EGV. EGV has seven subsites for sugar binding, termed A–G, aiding cleavage of the glycosidic bond between subsites D and E. As with lysozyme^{7,8,19} this cleavage is at the reducing end of the substrate. Cellobiose is shown bound to the 'leaving group' sites E and F. A water molecule, labelled 'wat', is bound near Asp 10. Modelling studies using hexasaccharide substrates indicate that, given a small movement (<1 Å), this water molecule would be suitably positioned to make a nucleophilic attack leading to inversion at C1. Crystals of the EGV–cellobiose complex were grown under the same conditions as the native enzyme (Table 1), with the addition of 5 mM cellobiose. The crystals are in space group $P2_1$ with cell dimensions $a=41.75$, $b=56.19$, $c=37.37$ Å and $\beta=93.40^\circ$. The structure was solved by molecular replacement using standard programs⁹. Data were collected to 1.9 Å using an RAXIS-II image-plate and a Cu rotating anode source and processed with the DENZO program (Z. Otwinowski, Yale University). Details of the data



This similarity between the lysozyme and EGV active site clefts is not surprising, given that these enzymes bind similar substrates, but it is not seen to this extent in other cellulolytic enzyme structures. Catalysis by lysozyme^{7–9} proceeds by an SN_1 mechanism with retention of configuration at C1. The oxycarbonium ion intermediate is stabilized by electrostatic interaction with Asp 52, steric factors dictating that the nucleophilic water molecule is only able to approach the sugar ring once the leaving group has departed. This, together with the participation of Glu 35 in directing the attacking water molecule, results in retention of configuration at C1. In contrast, catalysis by EGV, and by both CBH-II and CelD, proceeds with inversion (C. Schou, B. Henrissat & C. Gey, personal communication). The structure of a cellobiose complex, solved at 1.9 Å, reveals that the two catalytic aspartates are located such that a water molecule bound to Asp 10 could make nucleophilic attack without prior need for product departure (Fig. 3). This stereochemistry dictates that catalysis proceeds with inversion at C1. When cellobiose occupies the 'leaving group' site (subsites E and F), the loop between strands V and VI, disordered in the native structure, interacts with the cellobiose and is well defined.

The completely different folding in the three X-ray structures of cellulases is surprising because we might expect the differences in specificity between these cellulases to be achieved by changes to a basic framework, as is found in other enzyme families such as the ribonucleases, lipases and serine proteases. For those cellulolytic enzymes whose structures are known there is no identifiable relationship in the tertiary folding, but there is a pattern of convergence in their evolution towards similar catalytic sites. The insoluble nature of cellulose as a substrate presents many chemical and metabolic problems for the organism. This may be reflected by the presence of at least eight different families of cellulases and in their multi-domain character. The distinctive structure described here represents the first member of a new family of cellulases. Although the EGV structure is distinct as a cellulase, the presence of the same barrel topology as a polysaccharide-binding (chitin) plant defence protein strongly suggests there is a functional and evolutionary relationship between these two classes of protein.

Note added in proof: The endoglucanase V from *Humicola insolens* has recently been classified into cellulase family 'K' (ref. 20). □

quality and completeness are shown in Table 1. The map shown is an $F_o - F_c$ difference map, contoured at a level corresponding to about $0.33e \text{ Å}^{-3}$, from a model consisting of protein atoms only. The current crystallographic R-factor is 0.159, with r.m.s. deviations from stereochemical target values of 0.011 and 0.039 Å for bonds and angles, respectively.

- Strynadka, N. C. J. & James, M. N. G. *J. molec. Biol.* **220**, 401–424 (1991).
- Higashi, T. *J. appl. Crystallogr.* **22**, 9–18 (1989).
- CCP4, The SERC (UK) Collaborative Computing Project No. 4: *A Suite of Programs for Protein Crystallography* (Daresbury, UK, 1979).
- Otwinowski, Z. in *Proc. CCP4 Study Weekend* (eds Wolf, W., Evans, P. R. & Leslie, A. G. W.) 80–85 (CCP4, Daresbury, UK, 1991).
- Zhang, K. Y. J. & Main, P. *Acta Crystallogr.* **A46**, 377–381 (1990).
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta Crystallogr.* **A47**, 110–119 (1991).
- Brünger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
- Hendrickson, W. A. & Konnert, J. H. in *Structure, Conformation and Evolution* Vol. 1 (ed Srinivasan, R.) 43–57 (Pergamon, Oxford, 1981).
- Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).
- Richardson, J. *Adv. Prot. Chem.* **34**, 167–339 (1981).
- Bernstein, F. C. et al. *J. molec. Biol.* **112**, 535–542 (1977).
- Henrissat, B. & Bairoch, A. *Biochem. J.* **293**, 781–788 (1993).

ACKNOWLEDGEMENTS. We thank N. Sakabe for provision of data collection facilities at the Photon Factory synchrotron facility, H. Savage and T. Wilkinson for discussions, M. Hartshorn for assistance with the figures and A. Murzin (LMB, Cambridge) for pointing out the similarity between EGV and barwin. We thank the SERC and Novo-Nordisk a/s for financial support.

Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing

Melissa J. Moore & Phillip A. Sharp

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

EXCISION of introns from nuclear precursors to messenger RNAs (pre-mRNAs) by the spliceosome requires two distinct phosphodiester transfer (transesterification) reactions: exchange of a 3'–5' for a 2'–5' bond in the first step (lariat formation) and exchange of one 3'–5' phosphodiester for another in the second step (exon ligation)^{1–3}. We report here determination of the stereochemical course of each step using splicing substrates that contained a chiral phosphorothioate. This has provided strong evidence that both steps occur as single 'in-line' SN_2 nucleophilic displacement reactions, analogous to the mechanism of group I self-splicing introns^{4,5}. Additionally, because both steps are strongly inhibited by the R_P phosphorothioate diastereomer, but not by S_P , the spliceosome probably shifts between two active sites in catalysis of the two steps. Chemical and stereochemical similarities suggest that the catalytic site for the second step of spliceosomal processing is related to that of group I self-splicing introns.

Whether or not transesterification of a phosphodiester bond involves a covalent intermediate can be determined by following the stereochemical configuration around the reactive phosphorous: although a direct mechanism results in overall stereo-

Received 11 January; accepted 21 June 1993.

- Beguin, P. et al. *Biochem. Soc. Trans.* **20**, 42–46 (1992).
- Wood, T. M. *Biochem. Soc. Trans.* **20**, 46–53 (1992).
- Henrissat, B. *Biochem. J.* **280**, 309–316 (1991).
- Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J. K. C. & Jones, T. A. *Science* **249**, 380–386 (1990).
- Juy, M. et al. *Nature* **357**, 89–91 (1992).
- Ludvigsen, S. & Poulson, F. M. *Biochemistry* **31**, 8783–8789 (1992).
- Blake, C. C. F. et al. *Proc. R. Soc. B* **167**, 365–377 (1967).
- Phillips, D. C. *Proc. R. Acad. Sci.* **57**, 484–495 (1967).

chemical inversion, an indirect mechanism (involving a covalent intermediate) gives net retention^{6,7}. To follow the stereochemical course of each step of splicing, we synthesized pre-mRNA substrates (E1-IVS-E2) containing a single chiral phosphorothioate linkage at either the 5' or 3' splice site (legends to Figs 1, 2 and 3)⁸. A phosphorothioate is a phosphate analogue in which one of the two non-bridging oxygens has been replaced by sulphur to yield the R_P or S_P diastereomer (Fig. 4)⁷. The two diastereomers are distinguishable by their sensitivity to cleavage by particular enzymes. For example, ribonuclease T1 and snake venom phosphodiesterase cleave R_P phosphorothioate diesters much more readily than S_P ⁹⁻¹¹, whereas nuclease P1 cleaves the S_P diastereomer preferentially¹².

The spliceosome exhibits a marked stereoselectivity between the two phosphorothioate diastereomers (Fig. 1). Although neither step of splicing was appreciably affected by the S_P diastereomer, introduction of the R_P diastereomer at either splice site effectively blocked transesterification at that site. This is consistent with previous results from a random phosphorothioate substitution/interference assay showing that the R_P diastereomer interferes with splicing when present at either splice site¹³. The active S_P diastereomer, however, was not tested in the earlier study.

Differential inhibition by one phosphorothioate diastereomer presumably reflects stereospecific contacts between the substrate and active site constituents that are disrupted upon introduction of a sulphur atom⁷. For example, many phosphoryl transfer enzymes contain active site Mg^{2+} ions that are coordinated to one, but not the other, non-bridging oxygen on the reactive phosphate^{14,15}. Because Mg^{2+} does not readily accept sulphur as a ligand¹⁶, a differential reactivity with the two phosphorothioate diastereomers results.

Because the S_P diastereomer did not interfere with either step of splicing, this diastereomer was used to follow the stereochemical course of each reaction. For characterization of the first step, a splicing substrate (E1-IVS-E2 (I)) was constructed that

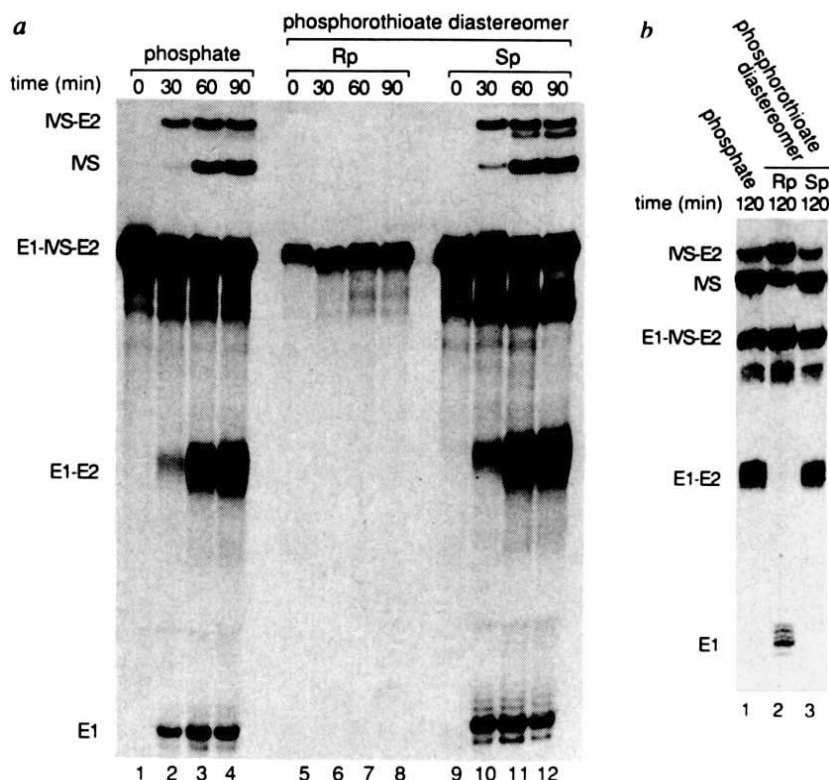
contained the S_P diastereomer at the 5' splice site and two radioactively labelled phosphodiester linkages (Fig. 2, upper). The radioactive phosphate adjacent to the 5' splice site served as a reporter for the stereochemical configuration of the phosphorothioate linkage in the E1-IVS-E2 (I) substrate, whereas that adjacent to the branch site adenosine served as a reporter for the configuration of the 2'-5' phosphorothioate linkage in the IVS (I) lariat product recovered after splicing.

Both E1-IVS-E2 (I) and IVS (I) were analysed by digestion with nuclease P1 and phosphodiesterase (Fig. 2, lower). Because nuclease P1 cannot cleave branched RNAs¹⁷, the bands in lanes 3 and 4 of Fig. 2 must represent the branched trinucleotides $^3A_3^2P(S)G_3^{OH}$ and $^3A_3^2P(S)G_3^{OH}$, respectively. Their electrophoretic mobility in this high percentage gel system is consistent with a charge to mass ratio intermediate to $^3G_{OH}$ and $^3G_P(S)G_{OH}$. The labelled phosphate in either trinucleotide was sensitive to removal with alkaline phosphatase (data not shown), confirming that the adenosine residue formed the core of the branch (and not the adjacent guanosine, for example). Therefore, the difference in migration between the phosphate and phosphorothioate-containing trinucleotides must reflect the replacement of a single non-bridging oxygen atom by sulphur.

Phosphodiesterase can cleave both 3'-5' and 2'-5' phosphodiester bonds¹⁷ and it cleaves R_P phosphorothioate linkages much more readily than S_P ^{10,11,18}. Thus, the 3'-5' S_P phosphorothioate linkage in the E1-IVS-E2 (I) substrate was resistant to phosphodiesterase cleavage (Fig. 2, lane 8). But the 2'-5' phosphorothioate linkage in the IVS (I) product could be cleaved by the enzyme (resulting in $^3A_{OH}$ rather than $^3A_2^2P(S)G_3^{OH}$ in lane 10), indicating that the 2'-5' linkage was in the R_P configuration. Therefore, the first step of pre-mRNA splicing proceeds with inversion of stereochemical configuration of the phosphate at the 5' splice site. This is strong evidence for a transesterification mechanism involving an odd number of 'in-line' SN_2 displacement reactions, with the simplest case being a single displacement with no covalent intermediate (Fig. 4, top left)^{6,7}.

FIG. 1 Splicing of pre-mRNA substrates containing a single R_P or S_P phosphorothioate at either the 5' or 3' splice site. **a**, Timecourse in min for splicing of E1-IVS-E2 RNAs containing a phosphate (lanes 1-4) or R_P (lanes 5-8) or S_P (lanes 9-12) phosphorothioate at the 5' splice site. **b**, Two-hour endpoint for splicing of E1-IVS-E2 RNAs containing a phosphate (lane 1) or R_P (lane 2) or S_P (lane 3) phosphorothioate at the 3' splice site. Inhibition of the first step of splicing by the R_P diastereomer at the 5' splice site is shown by the lack of splicing intermediates in part **a**, lanes 5-8, whereas the accumulation of splicing intermediates in part **b**, lane 2 indicates inhibition of the second step by an R_P diastereomer at the 3' splice site. The small amount of RNA migrating at the position of IVS product in **b**, lane 2 probably results from nuclease activity present in the nuclear extract as in ref. 8.

METHODS. GpG dinucleotides containing either a phosphate or phosphorothioate linkage were synthesized in the solid phase using a standard DNA synthesizer and off-line sulphurization^{34,35}. Following deprotection, dinucleotides were purified and diastereomers separated by C18 reverse-phase HPLC^{34,35}. Phosphorothioate stereochemistry and purity were determined by digestion of dinucleotides with ribonuclease T1, snake venom phosphodiesterase and nuclease P1 after phosphorylation with polynucleotide kinase and [γ -³²P]ATP. Incorporation of stereochemically pure dinucleotides at the 5' or 3' splice site of the HindIII T7-runoff transcript of plasmid pPIP8.5.A and subsequent splicing reactions were similar to our previous description⁸. Transcripts were radioactively labelled by inclusion of [α -³²P]UTP in transcription reactions.



The stereochemical course of the second step of splicing was determined similarly using a splicing substrate (E1-IVS-E2 (III)) that contained a S_P phosphorothioate linkage at the 3' splice site and a single radioactively labelled phosphate in E1 (Fig. 3, upper). The 3'-5' phosphorothioate linkage in the ligated exon product of splicing (E1-E2 (III)) displayed an opposite sensitivity to digestion by ribonuclease T1 and nuclease P1 as did that in an appropriately labelled substrate (E1-IVS-E2 (II)) (Fig. 3, lower). Therefore, the second step of nuclear pre-mRNA splicing also proceeds with complete inversion of stereochemistry, for which a single in-line S_N2 displacement reaction is the most likely mechanism (Fig 4, lower left)^{6,7}.

Because of the requirement for small nuclear RNAs in the spliceosome and structural similarities between the intermediates and products of pre-mRNA splicing and those of group II self-splicing introns, it has been proposed that the catalytic core of the spliceosome is composed largely of RNA^{19,20}. If so, what parallels can be drawn between the spliceosome and known RNA catalysts? Of the self-splicing introns, the group I intron from *Tetrahymena* 35S pre-ribosomal RNA is by far the best understood²¹. Excision of this intron proceeds through two steps, but without formation of a lariat intermediate (Fig. 4, right). Instead, the 3'-oxygen of exogenous guanosine is used as the nucleophile in the first step to bring about cleavage at the 5' splice site. In the second step, the 3' oxygen of the 5' exon (the leaving group in the first reaction) is the nucleophile that attacks at the 3' splice site to liberate the intron and join the two exons. Both steps proceed through a single in-line S_N2 transesterification mechanism^{4,5,22} and both are readily reversible.

For group I introns, there is good evidence that the two steps of splicing are catalysed as forward and reverse reactions within a single active site, in which the relative orientations of reactants is the same for the two steps. Both the free guanosine (the nucleophile in the first step) and the guanosine at the 3' splice

site (the leaving group in the second step) are linked to the reactive phosphate through a 3' oxygen and are bound in the same G-binding site within the intron^{23,24}. Additionally, the stereoselectivity for chiral phosphorothioate analogues by the *Tetrahymena* intron is consistent with the principle of microscopic reversibility. Because a direct S_N2 transesterification results in inversion of stereochemistry at the reactive phosphate, any phosphorothioate diastereomer preference in the first step of group I intron splicing (such as $R_P \gg S_P$) must be matched by an opposite preference in the second step (such as $S_P \gg R_P$) if the two steps are forward and reverse reactions mediated by a single active site. Thus although the R_P diastereomer has almost no effect on either the first step of group I intron splicing^{4,25} or the reverse of second step (intron insertion)²², this diastereomer is highly inhibitory for both the second step²² and the reverse of the first step⁵.

In contrast to the *Tetrahymena* intron, both steps of pre-mRNA splicing are strongly inhibited by R_P phosphorothioate linkages at the splice sites, and neither step is significantly affected by the S_P diastereomer. Therefore, the two steps of pre-mRNA splicing cannot be catalysed as forward and reverse reactions within a single active site. Neither is it likely that the two steps are catalysed as parallel reactions (both in the same direction) within a single active site, as it is difficult to imagine a single site that could direct stereoselective synthesis of both a 2'-5' branched RNA in the first step and a 3'-5' linear RNA in the second step. Rather, it seems more likely that, in progression from the first step to the second step, there is an interconversion between two active sites within the spliceosome. This does not necessitate that the two active sites be entirely different; it is possible that they are overlapping, with each containing a subset of components specific to that site. In fact, a recently demonstrated interaction between the first and last nucleotides of a pre-mRNA intron in *Saccharomyces cerevisiae*²⁶ may indicate that

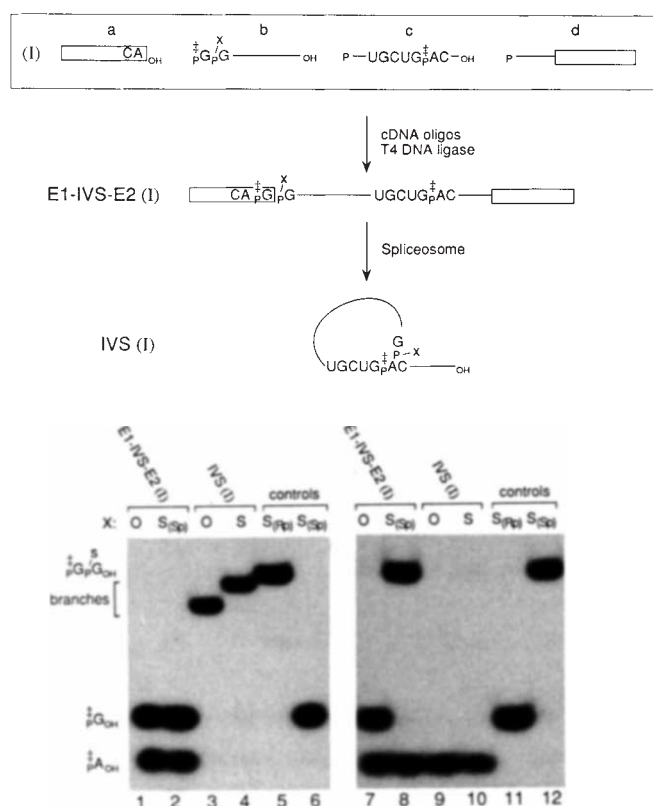


FIG. 2 Determination of the stereochemical course of lariat formation. Upper, Synthesis of E1-IVS-E2 substrates containing two ^{32}P -labelled phosphates ($\dagger$) and either a phosphate ($X=O$) or S_P phosphorothioate ($X=S$) linkage at the 5' splice site. E1-IVS-E2 RNAs were made initially in four segments (a, b, c and d) which were ligated in a single reaction (I), to give an overall 4% yield of full-length E1-IVS-E2 (I). Lariat product RNAs, IVS (I), were generated by incubation of the E1-IVS-E2 (I) substrates (1×10^6 c.p.m.) under splicing conditions⁸ for 120 min, and then purified on a 15% polyacrylamide gel. Lower, Digestion of E1-IVS-E2 (I) substrates and IVS (I) products with nuclease P1 (lanes 1-6) and phosphodiesterase (lanes 7-12). Both enzymes cleave all 3'-5' phosphodiesterates to yield nucleotide 5'-monophosphates. But nuclease P1 cleaves neither branched RNAs (lanes 3 and 4; ref. 17) nor R_P phosphorothioates (lane 5; ref. 12), and phosphodiesterase does not readily cleave S_P phosphorothioates (lane 12; refs 10, 11). METHODS. RNA a was identical to RNA(1-54) in ref. 8. RNA b, corresponding to pPIP85.A nucleotides 55-145⁸, was transcribed with either GpG or GpG(S_P) as a primer, and subsequently phosphorylated with polynucleotide kinase and [γ - ^{32}P]ATP. RNA c, 5'-pGGGUGCUGP_XACU-3', was generated by oligonucleotide transcription³⁶ in the presence of GMP and [α - ^{32}P]ATP. RNA d (C. Query, M.J.M. & P.A.S., manuscript in preparation) was also initiated with GMP. RNA ligation reactions were done with T4 DNA ligase as described⁵. Segments a and b were bridged by cDNA(68-44)⁸; a second bridge of 31 nucleotides spanned segments b, c and d (C. Query, M.J.M. & P.A.S., manuscript in preparation). Control RNAs (lanes 5, 6, 11 and 12) were generated by joining segments a and b only, where segment b was initiated with either GpG(R_P) or GpG(S_P) and phosphorylated with [γ - ^{32}P]ATP. Nuclease P1 (0.3 μ g enzyme) and phosphodiesterase (0.6 μ g enzyme) digestions were done in 3 μ l 30 mM sodium acetate pH 5.2, 0.1 mg ml⁻¹ *E. coli* transfer RNA or 50 mM Tris-HCl pH 7.5, 0.1 mM EDTA, 10 mM magnesium chloride, 0.1 mg ml⁻¹ *E. coli* tRNA, respectively, for 20 min at 37 °C. Digestion products were separated by electrophoresis through a 25% (19:1 crosslinking) polyacrylamide gel as described⁸.

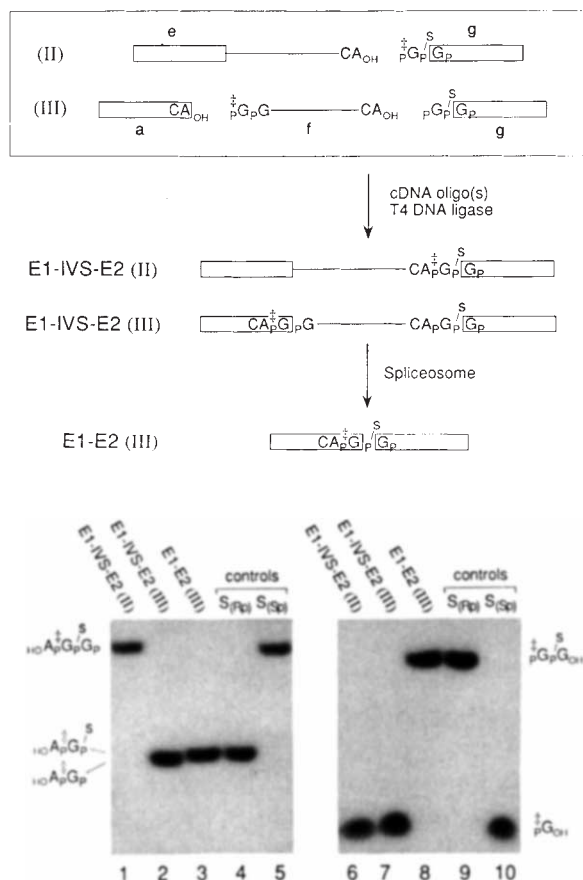


FIG. 3 Determination of the stereochemical course of exon ligation. Upper, Scheme for synthesis of E1-IVS-E2 substrates containing a S_P phosphorothioate linkage at the 3' splice site and a single [^{32}P]-labelled phosphate ($\ddagger$) adjacent to either the 3' (E1-IVS-E2 (II)) or 5' (E1-IVS-E2 (III)) splice site. E1-E2 (III) was generated by incubation of the E1-IVS-E2 (III) substrate (1.5×10^6 c.p.m.) under splicing conditions⁸ for 90 min, and then purified on a 15% polyacrylamide gel. Lower, Digestion of E1-IVS-E2 (II), E1-IVS-E2 (III) and E1-E2 (III) with ribonucleases A and T1 (lanes 1–5) and nuclease P1 (lanes 6–10). Ribonuclease A cleaves after pyrimidines, whereas T1 cleaves only after guanosine; both enzymes generate oligonucleotides with 3'-monophosphates and ribonuclease T1 cleaves R_P phosphorothioates preferentially over S_P (lanes 4 and 5; ref. 9). Nuclease P1 generates 5'-phosphate mononucleotides and cleaves S_P phosphorothioates preferentially over R_P (lanes 9 and 10; ref. 12).

METHODS. RNA segments a, e and g were identical to RNA(1–54), RNA(1–182) and RNA(183–236), respectively, in ref. 8, except that segment g was initiated with $Gp^S G(S_P)$ and phosphorylated with either [γ - ^{32}P]ATP (for ligation II) or unlabelled ATP (for ligation III). RNA f corresponded to pPIP85.A nucleotides 55–182. RNA ligation reactions and bridging cDNAs were similar to those in ref. 8. Ribonuclease digestions were done in 4 μ l 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.1 mg ml⁻¹ *E. coli* tRNA with 0.4 μ g ribonuclease A and 1 unit of ribonuclease T1 (>350 units μ g⁻¹). Nuclease P1 digestions were as in Fig. 2, except in a volume of 4 μ l.

the product of the first step, the lariat intermediate, constitutes part of the active site for the second step by providing a binding site for the AG dinucleotide at the 3' splice site.

Additional evidence supports a two active site model for the spliceosome¹. Nucleotide changes at branch and splice sites^{27–29}, temperature-sensitive mutations in splicing factors^{2,3,30}, modulation of the phosphorylation state of factors³¹ and mutations in U2 and U6 snRNAs^{32,33} can all arrest splicing after the first step. At a more chemical level, introduction of a 2' modified ribose adjacent to the reactive phosphate at either splice site produced

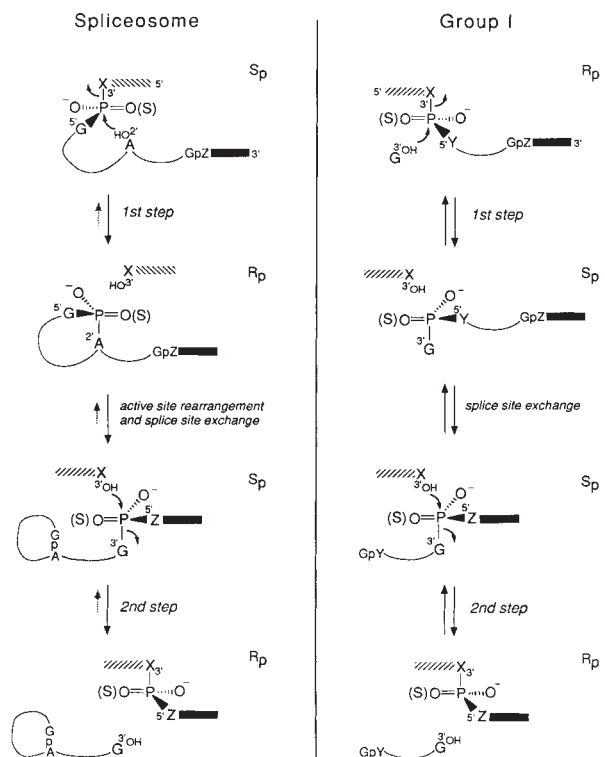


FIG. 4 Chemical mechanisms and stereochemical configurations of the two steps of pre-mRNA splicing (left) and group I intron self-splicing (right). First-step differences between spliceosomal and group I splicing include opposite phosphorothioate diastereomer preferences, different reactive nucleophiles (2'-OH versus 3'-OH), and differential placement (5' versus 3') of the conserved guanosine (G) with respect to the phosphate at the 5' splice site. Second-step similarities include preference for the S_P phosphorothioate diastereomer, use of a 3'-OH as the nucleophile and a conserved guanosine (G) attached through a 3' oxygen to the phosphate at the 3' splice site. The non-bridging position (S) at which sulphur is not significantly inhibitory, along with the corresponding phosphorothioate diastereomer designation (R_P or S_P), is indicated for the substrate for each step. The diastereomer that results is indicated for the product of each step. (For the group I case, the stereochemical outcome for the second step has not been directly determined, but is inferred from studies of a ligation reaction thought to be analogous to exon ligation⁵ and of the reverse of exon ligation²².) Small dashed arrows indicate the potential (but not yet observed) reversibility of each step of nuclear pre-mRNA splicing. X, Y and Z indicate nucleotide positions that are not highly conserved.

markedly different effects on the two steps of splicing. This suggests differential recognition of these riboses in each step⁸.

If the two steps of nuclear pre-mRNA splicing are RNA catalysed, it is interesting to speculate that the two active sites may be homologous to different known RNA catalysts. The first step may reflect the activity of a catalytic site common to pre-mRNA splicing and group II introns. The characteristics of such a site would include recognition of a conserved guanosine linked to the reactive phosphate via a 5' oxygen, nucleophilic activation of a 2' oxygen on a nucleoside bulged from a duplex segment, and the ability to accommodate an S_P phosphorothioate diastereomer in the forward reaction. The spliceosomal active site for the second step might be similar to that of group I introns. The shared characteristics of this site would include recognition of a conserved guanosine linked to the reactive phosphate at the 3' position, nucleophilic activation of a 3' oxygen at the terminus of an RNA segment, and accommodation of an S_P phosphorothioate diastereomer in the forward reaction (as defined by the

second step reaction of group I introns). We propose that the spliceosome generates a group I-like catalytic site to execute the second step of pre-mRNA splicing. □

Received 1 June; accepted 12 July 1993.

- Moore, M. J., Query, C. C. & Sharp, P. A. In *The RNA World* (eds R. Gesteland & J. Atkins), 303–357 (Cold Spring Harbor Laboratory Press, New York, 1993).
- Guthrie, C. *Science* **253**, 157–163 (1991).
- Ruby, S. W. & Abelson, J. *Trends Genet.* **7**, 79–85 (1991).
- McSwiggen, J. A. & Cech, T. R. *Science* **244**, 679–683 (1989).
- Rajagopal, J., Doudna, J. A. & Szostak, J. *Science* **244**, 692–694 (1989).
- Knowles, J. R. A. *Rev. Biochem.* **49**, 877–919 (1980).
- Eckstein, F. A. *Rev. Biochem.* **54**, 367–402 (1985).
- Moore, M. J. & Sharp, P. A. *Science* **256**, 992–997 (1992).
- Eckstein, F., Schulz, H. H., Rüterjans, H., Haar, W. & Maurer, W. *Biochemistry* **11**, 3507–3512 (1972).
- Burgers, P. M. J. & Eckstein, F. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4798–4800 (1978).
- Bryant, F. R. & Benkovic, S. J. *Biochemistry* **18**, 2825–2828 (1979).
- Potter, B. V. L., Connolly, B. A. & Eckstein, F. *Biochemistry* **22**, 1369–1377 (1983).
- Maschhoff, K. L. & Padgett, R. A. *Nucleic Acids Res.* **20**, 1949–1957 (1992).
- Yarus, M. *FASEB J.* **7**, 31–39 (1993).
- Steitz, T. A. & Steitz, J. A. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6498–6502 (1993).
- Pecoraro, V. L., Hermes, J. D. & Cleland, W. W. *Biochemistry* **23**, 5262–5271 (1984).
- Reilly, D. J., Wallace, J. C., Melham, R. F., Kopp, D. W. & Edmonds, M. *Meth. Enzym.* **180**, 177–191 (1989).
- Nelson, P. S., Bach, C. T. & Verheyden, J. P. H. *J. org. Chem.* **49**, 2314–2317 (1984).
- Sharp, P. A. *Cell* **42**, 397–400 (1985).
- Cech, T. R. *Cell* **44**, 207–210 (1986).
- Cech, T. R. A. *Rev. Biochem.* **59**, 543–568 (1990).
- Suh, E.-R. & Waring, R. B. *Nucleic Acids Res.* **20**, 6303–6309 (1992).
- Michel, F., Hanna, M., Green, R., Bartel, D. P. & Szostak, J. W. *Nature* **342**, 391–395 (1989).
- Been, M. D. & Perrotta, A. R. *Science* **252**, 434–437 (1991).
- Herschlag, D., Piccirilli, J. A. & Cech, T. R. *Biochemistry* **30**, 4844–4854 (1991).
- Parker, R. & Siliciano, P. G. *Nature* **361**, 660–662 (1993).
- Hornig, H., Aebi, M. & Weissmann, C. *Nature* **324**, 589–591 (1986).
- Aebi, M., Hornig, H. & Weissmann, C. *Cell* **50**, 237–246 (1987).
- Lamond, A. I., Konarska, M. M., Sharp, P. A. *Genes Dev.* **1**, 532–543 (1987).
- Frank, D., Patterson, B. & Guthrie, C. *Molec. cell. Biol.* **12**, 5197–5205 (1992).
- Mermoud, J. E., Cohen, P. & Lamond, A. I. *Nucleic Acids Res.* **20**, 5263–5269 (1992).
- Fabrizio, P. & Abelson, J. *Science* **250**, 404–409 (1990).
- McPheeters, D. S. & Abelson, J. *Cell* **71**, 819–831 (1992).
- Stec, W. J., Zon, G., Egan, W. & Stec, B. J. *Am. chem. Soc.* **106**, 6077–6079 (1984).
- Applied Biosystems Model 380 User Bulletin 44, 1–18 (1987).
- Milligan, J. F. & Uhlenbeck, O. C. *Meth. Enzym.* **180**, 51–62 (1989).

ACKNOWLEDGEMENTS. We thank R. Cook, H. LeBlanc and S. Schultz of the MIT Biopolymers Facility for synthesizing and purifying the GpG dinucleotides; and members of the Sharp laboratory (B. Blencowe, J. Crispino, T. Kristie, A. McMillan, R. Meyers and C. Query) for critical reading of the manuscript. This work was supported by Helen Hay Whitney and American Cancer Society, Mass. Division, Inc., postdoctoral fellowships (M.J.M.), a grant from NIH (P.A.S.) and partially by a NCI Center core grant.

ERRATA

Crystal structure of active elongation factor Tu reveals major domain rearrangements

Harald Berchtold, Ludmila Reshetnikova, Christian O. A. Relser, Norbert K. Schirmer, Mathias Sprinzl & Rolf Hilgenfeld

Nature **365**, 126–132 (1993).

COPIES of the 9 September issue printed in Europe contained an error in this article. On page 127, instead of Fig. 2a, the image for Fig. 3a was repeated. US and Japanese editions were correct. Figure 2a is shown below. □

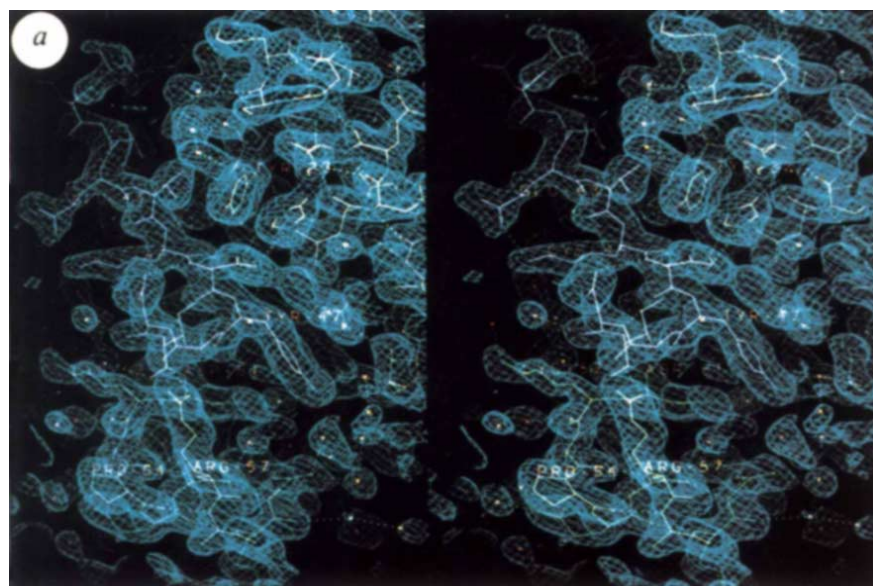
Vegetation effects on the isotope composition of oxygen in atmospheric CO₂

Graham D. Farquhar, Jon Lloyd, John A. Taylor, Lawrence B. Flanagan, James P. Syvertsen, Kerry T. Hubick, S. Chin Wong & James R. Ehleringer

Nature **363**, 439–443 (1993)

SUBSCRIBERS to *Nature* in the United States and Canada will have received an incorrect version of this letter, in which colour Figs 2 and 3 have been transposed. In copies of *Nature* received by the rest of the world, the figures are correct as shown.

Also, the symbol δ_e used in the first sentence of the second new paragraph on page 442 should be δ_r , the oxygen isotope composition of respired CO₂. □



Systematic identification of clone coordinates from high-density filters

C. Peter Jones

A system is described for accurately translating the positive signal, obtained after hybridization of high-density colony filters, into clone coordinates. Calibration marks are made on the filter at the same time and in register with the clones. A combination of digitizing tablet and software allows for the retrieval of clone coordinates directly from the autoradiograph.

With the advent of current screening strategies based around the 96-well format microtitre plate, various groups have incorporated the use of high-density screening systems whereby the area usually occupied by one 96-well format has up to 36 off-sets placed in the equivalent area¹⁻⁸. This has been based on arrays of either 2×2 , 3×3 , 4×4 , 5×5 or 6×6 , resulting in colony densities between 384 to 3,456 in an area of approximately 80×120 mm. This condensation of clones allows for a considerable economy in the number of filters that have to be screened.

When dealing with regular arrayed high-density filters, the retrieval of the correct coordinates from a positive hybridization signal can be difficult. This problem is compounded if there is very little background hybridization on the filter. Previous methods have relied on hybridizing the filter with either ³⁵S-prepared probes¹ of vector, aligning an acetate sheet (with the coordinates for each position marked on) with the original filter and the autoradiograph on a light box until all three are in register or using radioactive ink to provide the registration points⁷. These methods can be time consuming when dealing with more than a few filters, and can be prone to error, especially on the higher density filters^{9,10}.

Spotting in register

To improve the accuracy and speed of colony localization-to-hybridization signal, a system was designed that automatically reports the positive hybridization signal as well as a

position in the microtitre plate originally used to make the filter. The key feature of the system is the ability to place the colonies at high density on to the membrane in register with the pins of the spotting device. Currently, a 384-pin spotting tool (Genetix Ltd, Christchurch, UK) is used with 15-mm long location pins positioned 4 mm from positions A1, A24 and P24 (Fig. 1). The pins protrude 1 mm beyond the end of the spotting tool. In order for the spotting tool to work, the 384-well Quad plate⁷ has holes in the plate outside the wells such that the pins on the spotting tool can reach the bottom of the wells (Fig. 1). The location holes also offer the possibility of using them as anchor points for gripping tools when moving plates from one location to another. To transfer cells, the spotting tool is placed in the wells of the Quad plate and then placed onto the membrane. The location pins make small marks in the membrane, exactly in register with the spotting tool pins. The membrane is then processed by conventional means to fix the DNA from the cells to the membrane¹¹⁻¹². Before the membrane is hybridized, a small quantity of 'Cats Eyes' fluorescent ink is applied to each location pin mark. The ink permanently attaches itself to the membrane and is not removed by hybridization or stripping protocols. The filter is then hybridized by conventional means¹², however, just before the filter is put down, the Cats Eyes are activated by brief exposure to ultraviolet light on a transilluminator. After radiography, the Cats Eyes can be readily

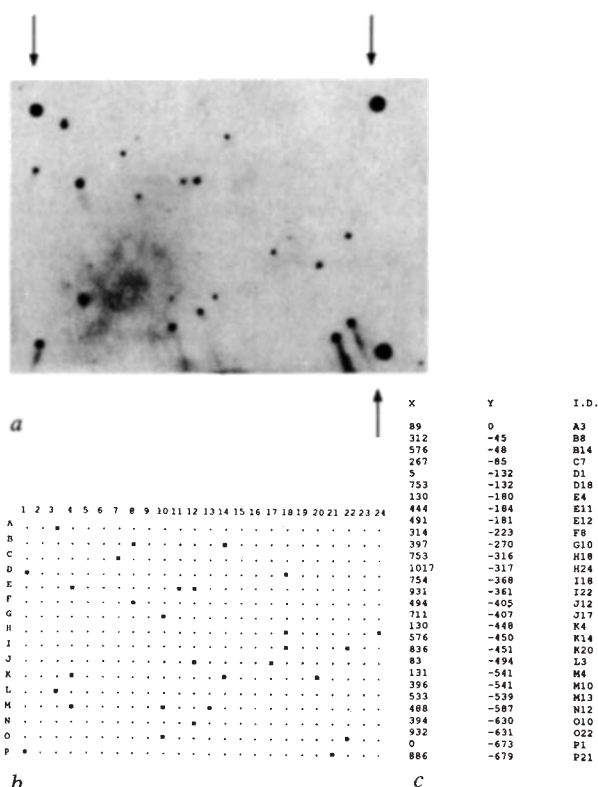


FIG. 2a, An autoradiograph after hybridization of cDNAs³ with a complex probe from a chromosome paint of chromosome 9 showing the positions of the Cats Eyes calibration points (arrows) and the hybridization signals. b, Graphical output of the AutoRad program showing the positive signals as dark spots. c, Clone identification coordinates for the positive clones.

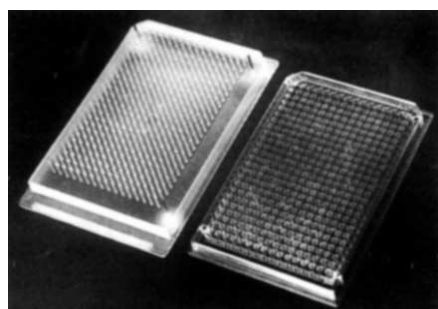


FIG. 1 384-well Quad plate (right) showing the position of the location holes and 384 spotting tool (left) showing the location pins.

seen on the autoradiograph along with any positive hybridization signals (Fig. 2a).

Data retrieval

Because the Cats Eyes are in register with the spotting pins, the location of the positive hybridization signals can be retrieved either by the use of an acetate overlay with the reference points marked on or by the use of a digitizing tablet coupled to an Apple Macintosh computer running the AutoRad software package. In this system, quantitative and qualitative information about the hybridization signal is determined by the user.

The combination of software and digitizing tablet enables the system to retrieve

hybridization coordinates from all clone densities based on either the microtitre plate or Quad-plate format. The system is calibrated by clicking the cursor of the digitizing tablet over each Cats Eye in turn. After calibration, clone location is obtained by placing the cursor over the hybridization signal and clicking. The software then translates the hybridization signal coordinate into clone location and displays the clone identification on screen (Fig. 2c). The software takes into account the shrink and swell of the filter and assigns the digitized reference to the nearest position expected from the spotting tool. By way of a double check, the colony coordinates can be converted into a graphical form which can be printed (Fig. 2b). Collected coordinates can be stored and manipulated in standard Macintosh files. It will also be possible to export the clone coordinate files to robotic devices for automatic clone retrieval either from storage plates or colony-containing plates¹⁻⁸.

The above system has speeded up the location of positive hybridization signals from high-density filters especially on low background examples or where complicated hybridization patterns are obtained. The accuracy of the system has been determined as ± 0.1 mm. However, in practice, this is effectively ± 0.3 mm (still well within acceptable limits for the localization of clones arrayed at high density) because of the constraints imposed by manual calibration of the system. □

C. Peter Jones is at the University of Cambridge, Human Molecular Genetics Research Group, Department of Pathology, Tennis Court Road, Cambridge CB2 1QP, UK. C. P. J. was supported by funds provided by the MRC for the original version of AutoRad software based on a sonic digitizer, commissioned from Visionways (Ringwood, Hampshire, UK), written and further developed by Martin Davies. For more information, fill in reader service number 100.

1. Nizetic, D. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 3233-3237 (1991).
2. Uber, D. C., Jaklevic, J. M., Theil, E. H., Lishanskaya, A. & McNeely, M. R. *BioTechniques* **11**, 642-646 (1991).
3. Jones, P., Watson, A., Davies, M. & Stubbings, S. *Nucleic Acid Res.* **20**, 4599-4606 (1992).
4. Bentley, D. R. et al. *Genomics* **12**, 534-541 (1992).
5. Medvik, P. A., Hollen, R. M., Roberts, R. S., Trimmer, D. & Beugelsdijk, T. J. *Int. J. Genomic Res.* **1**, 17-23 (1992).
6. Ross, M. T. et al. in *Techniques for the Analysis of Complex Genomes* (ed. Anand, R.) 137-154 (Academic, New York, 1992).
7. Meier-Ewert, S., Maier, E., Ahmadi, A., Curtis, J. & Lehrach, H. *Nature* **361**, 375-376 (1993).
8. Watson, A., Smaldon, N., Lucke, R. & Hawkins, T. *Nature* **362**, 569-570 (1993).
9. Hoheisel, J. et al. *Cell* **73**, 109-120 (1993).
10. Mizukami, T. et al. *Cell* **73**, 121-132 (1993).
11. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961-3965 (1975).
12. Berger, S. L. & Wahl, G. M. in *Guide to Molecular Cloning Techniques* (eds Berger, S. L. & Kimmel, A. R.) 415-423 (Academic, San Diego, 1987).

Scientific supplies

This week's selection includes a plant photosynthesis meter, a genotoxicity testing system and a digital imaging spectrophotometer.

THE EARS-PPM plant photosynthesis meter available from Cottenham Instruments is a chlorophyll fluorometer designed for the rapid determination of the photosynthetic light-use efficiency in green plants (*Reader Service No. 101*). Unlike conventional plant fluorometers, which usually measure the Kautsky induction curve on dark-adapted leaves (a time-consuming procedure that is not suitable for fast assessment and practical application), the PPM measures the plant in its ambient light. In the basic mode, light-use efficiency is displayed right away, so series of measurements can be carried out quickly. The ambient photosynthetic active radiation is measured simultaneously, enabling the calculation of gross photosynthesis. In another mode, PPM measures chlorophyll fluorescence. By means of analog output or PC control, it is possible to record fluorescence transients such as the Kautsky induction curve. Weighing only 1 kg, the PPM is particularly suited for taking measurements in the field.

A practical 'overnight test' for genotoxicity is now being offered by Microbics (*Reader Service No. 102*). The Mutatox test system measures genetic damage to test organisms caused by agents in a test sample — information that is of particular interest because many genotoxic agents are mutagenic



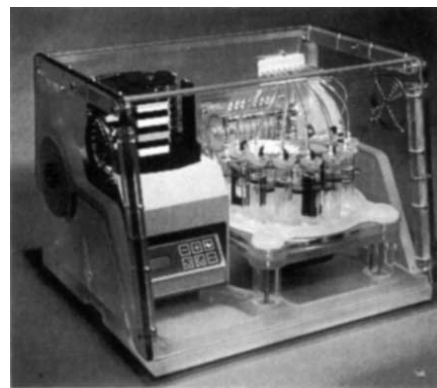
Mutatox testing system for genotoxic agents — see Microbics.

and/or carcinogenic. The system uses a dark strain of normally luminescent microorganisms as test specimens. If some agent in the test sample alters genetic control of the light-producing mechanism in the organisms, the light switches on. The light output of the organisms is measured 16 hours after the organisms have been exposed to various concentrations of the materials being tested. Microbics says that a test sample is considered genotoxic if the amount of light produced by the cells after exposure to it is at least twice the light produced by a non-genotoxic control sample. Samples may be

tested with or without a metabolic activator provided with each test kit. The Mutatox test can be performed with the modified Microbics Microtox toxicity test equipment; expensive special laboratory facilities are therefore unnecessary.

Protein purification systems

The Hoefer IsoPrime multi-chambered electrofusing unit can be used to purify proteins with high resolution based on their isoelectric points (*Reader Service No. 103*).



Process up to several hundred milligrams of protein per run with Hoefer's IsoPrime.

Hoefer says that the unit can process up to several hundred milligrams of protein per run. Proteins are separated by isoelectric focusing in solution free of carrier ampholytes, acrylamide and its byproducts, and other matrix materials. Each protein is sequestered in a liquid-filled chamber bounded by isoelectric membranes of pH values above and below the pI value of the protein. By careful selection of the pH of each membrane, a protein can be isolated in a different chamber from contaminating species. The separation unit is assembled with up to eight chambers separated by membranes of user-defined pH values. Software is used to calculate the acrylamido buffer composition for each membrane. The system includes an eight-channel peristaltic pump for circulating samples in each chamber through corresponding 30 ml reservoir vials.

Gilson Medical Electronics introduces ProTech, a low-pressure liquid chromatography system for biochromatography (*Reader Service No. 104*). The system is designed to offer a flexible and reliable alternative to more sophisticated systems for purifying biologically active proteins. ProTech automatically generates any composition gradients required by ion exchange, hydrophobic interaction, affinity and size-exclusion chromatography, and is designed



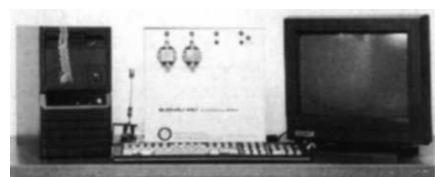
ProTech — the new low-pressure system from Gilson Medical Electronics.

to purify from several milligrams up to 100 g of protein. The system includes two Minipuls 3 single-channel peristaltic pumps for creating binary gradients. Each pump has ten rollers which generate a smooth, pulse-free flow. The flow-rate range of the Minipuls 3 is 0.05–50 ml min⁻¹ with a standard pump head for laboratory separations; up to 220 ml min⁻¹ with a high flow pump head for pilot-scale purifications. The ability of the Minipuls 3 to operate at 70 p.s.i. allows ProTech to use smaller, more efficient, high-resolution LC columns for higher purity and optimum yields. The system's fraction collector collects by peak, time, drop, or manual, with 10 collection windows in any mode. More than 20 different racks, including three that are thermostatic, are available. ProTech can also be configured for fixed wavelength UV/visible detection at 254 nm or 280 nm, with additional lamp/filter assemblies available, or for variable wavelength UV and visible detection, monitoring wavelengths between 190 nm and 700 nm. Control of the system is by IBM-compatible software.

■ Chromatography

During the process of enantioseparation by HPLC, impurities can cause contamination of the HPLC column. E. Merck has come up with a new set of **guard cartridges**, LiChroCart 4-4 ChiraDex and ChiraDex Gamma, to protect valuable HPLC cartridges (LiChroCart 250-4 ChiraDex or ChiraDex Gamma) against contamination (*Reader Service No. 105*). Guard cartridges are available in an economical 10-cartridge package.

Micro-Tech Scientific has introduced a new **micro-liquid chromatography system** that is intended for applications that include LC-MS, catecholamine and protein

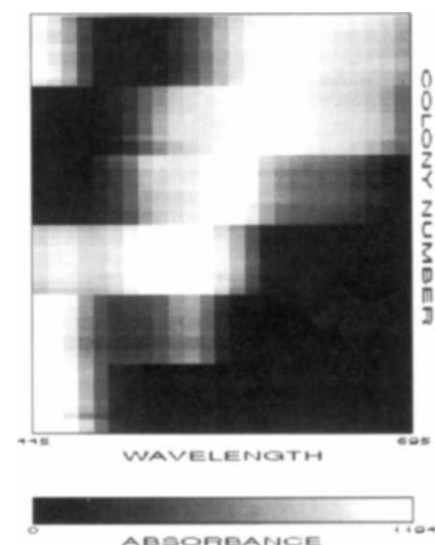


Micro-Tech's micro-LC system.

analysis (*Reader Service No. 106*). The Prodigix-4P pump is designed to generate pulse-free flows to less than 1.0 µl min⁻¹ and reproducible gradients to less than 10 µl min⁻¹. This constant-flow reciprocating pump is designed for packed capillary and microbore chromatography. It can be used independently or in a micro-optimized LC system, which includes a UV detector, injector, PC data system and microcolumn.

■ Analytical instrumentation

ColonyImager from Kairos is a turnkey **digital imaging spectrophotometer** that combines the wavelength resolving power of a spectrophotometer with the spatial resolution of a CCD camera (*Reader Service No. 107*). Kairos says that, in minutes, the ColonyImager can measure the spectra of hundreds of colonies directly from the surface of the Petri dish. An interactive user interface facilitates fully automated spectral acquisition, processing and display. Stated applications include natural products chemistry and diagnostics that use either intrinsic



ColonyImager from Kairos.

pigments or extrinsic colorimetric indicators. ColonyImager is radiometrically calibrated, yielding absorption spectra, reflectance spectra, or fluorescence emission images from 400 nm to 950 nm. Microtitre trays can also be imaged from the blue to near-infrared at 2 nm resolution. The system includes CyberDIS software which provides tools for rapidly sorting, comparing and visualizing hundreds of spectra under Microsoft Windows. The contour image shown above uses a grey-scale map (or pseudocolouring scheme) to display the absorption spectrum of every feature in a target (Petri dish). Each row represents a single feature (colony) and each column encodes absorption data as a function of wavelength. Features with common spectral properties can be grouped by a variety of sorting algorithms. In this example, Kairos says that six different spectral categories were detected.

Interactive windows in CyberDIS enable the researcher to correlate specific spectra with the position of features within the target and to display conventionally formatted absorption spectra (absorbance versus wavelength) for any feature.

Finnigan-MAT has two new **mass spectrometers**: a triple-stage quadrupole mass spectrometer, the TSQ 7000, and a single-stage quadrupole, the SSQ 7000 (*Reader Service No. 108*). The 7000 series has been designed with longer hyperbolic quadrupole



Finnigan-MAT's TSQ 7000 triple-stage quadrupole mass spectrometer.

mass analysers, offering higher resolution and a new detection system designed to improve sensitivity and reduce noise. A new hybrid turbomolecular pumping system gives the new instruments faster pumpdown and low baseline operating pressures for higher sensitivity. The instruments accommodate a complete range of sample inlets and ionization techniques. Three inlet ports at the ion source allow attachment of multiple interfaces to the same system and easy switching among them. The sample inlets and ionization techniques include gas chromatograph with electrospray ionization and positive or negative chemical ionization, direct insertion probes (liquid-cooled solids probe, desorption chemical ionization probe, fast atom bombardment and BioProbe), and liquid chromatography interfaces (atmospheric pressure ionization, including electrospray and atmospheric pressure chemical ionization, particle beam interface, thermospray and supercritical fluid chromatograph). Four serial and five analog input ports allow all auxiliary instruments to be operated under data system control and the data to be processed in real time. Laboratory equipment such as HPLC pumps, GCs and autosamplers can be connected to the data system through standard RS-232 ports. Additionally, five independent analog inputs, two analog outputs and three TTL input/output lines provide simultaneous control and acquisition of many different experimental parameters such as analog signals from a UV detector.

■ Microscopes/image analysis

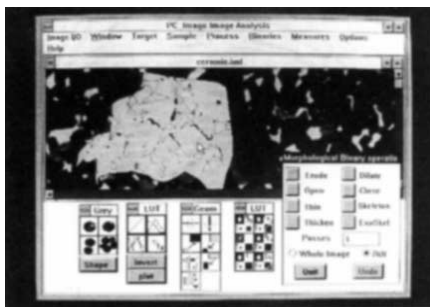
Carl Zeiss says that its new Fluar series of **microscope objectives designed for cellular calcium measurements** provide high light transmission by virtue of their high

numerical aperture and reduced number of glass elements (*Reader Service No. 109*). Retaining all the advantages of the infinity colour-corrected system optical design, the Fluor objectives have a stated light transmission at 340 nm of greater than 70 per cent, making them useful for epifluorescence measurements of intracellular calcium when FURA 2 is used as a calcium indicator. The objectives are also said to be suitable for transmitted-light brightfield, phase contrast and differential interference techniques. This set-up is useful for viewing and documenting normal morphological features of the fluorochrome-loaded cells. Fluor objectives are recommended for use with Axiovert, Axioskop and Axioplan microscopes, and the LSM 410 inverted confocal laser scanning microscope. Three types are available: the Fluor $\times 10/0.5$, $\times 40/1.3$ oil and $\times 100/1.3$ oil.

The new 1910 **field-emission scanning electron microscope (SEM)** from Amray, intended for analytical laboratories, materials and biological research facilities, and electronics/semiconductor production and quality assurance operations, features the ability to accommodate specimens up to 6 inches in diameter, a resolution of 1.5 nm at 30 kV and several operational enhancements (*Reader Service No. 110*). At the heart of the 1910 FE SEM is Amray's three-element Model 305 Schottky field-emission gun, which allows the kV to be easily changed throughout the instrument's 0–30-kV range, with few column adjustments, while the image remains clear and constant within the field of view. Amray says that as these field-emission guns are shipped under vacuum, with factory installed, fully stabilized and characterized tips, source changes can be made in a matter of hours. In addition, the company says that no laborious, time-consuming maintenance procedures are required to maintain a vacuum in the 10^{-10} torr range. The three-stage pumping system is also designed to ensure trouble-free operation and to eliminate any possibility of gun contamination from samples prone to outgassing. Other design features include a five axes, 4 inch motion universal specimen stage; automatic image focus, contrast and brightness; digital image processing; and

computer control and image archiving with the Power Windows SEM control package. In addition, the 1910 FE is also energy dispersive X-ray compatible.

A new system for true colour image analysis is now available from Synoptics (*Reader Service No. 111*). ColourPC_Image is a Windows-based package that operates with the Sprynt framestore for PC-compatible computers. This general purpose image analysis and measurement program uses the flexibility of Windows to simplify manipulation of the RGB colour image components. Image processing and image sampling functions can be performed on all three of the component images or on just one or two of the components. The Sprynt framestore uses the Intel i860 fast parallel processing chip, which accelerates all of the image processing and analysis functions available. Synoptics says that typical examples for the system are 0.54 s for a 512×512 rotate 30, less than 0.15 s for a linear filter on



Colour Image analysis from Synoptics.

a 512×512 image and 2.6 s for a 256×256 FFT (copy to real, clear imaginary, FFT-2D). In addition, Sprynt has output look-up tables that allow interactive thresholding to be performed in both monochrome and colour. Analysis and processing functions include arithmetic, geometric, convolution and non-linear processing, grey level morphology and fourier processing. Primary field and object measurements can be made, while derived measurements can be produced using the expression generator. Data can be displayed as listings or further analysed to produce histograms, graphs or scattergrams. Other features include point sampling, line profiles and three-dimensional plots of grey level distribution. Data may be exported to other Windows programs. ColourPC_Image can be adapted for more specialized applications by the addition of extra functions using the PC_ImXtnd facility.

VoxBlast, which is being distributed by VayTek, is a software package designed to provide accurate **measurements and visualization of volume data sets** (*Reader Service No. 112*). The program runs on Macintosh (series II), IBM AT (Windows) and UNIX (SGI, Sun, HP, IBM Risc 6000) computers; the UNIX program is written in X11 and Motif, making it a single copy/multiple



Three-dimensional measurements for volume data sets — see VayTek.

user version. VoxBlast has both a voxel-based and a polygon-based renderer, and can extract 2 1/2D or 3D surfaces from the voxel data set and merge the voxel data set and the polygon data set together again. The program coordinates the 2D slicing window and the 3D projection window for measurements. The user can place a cutting plane at any angle and at any depth in the volume. The cross section is shown in the 2D image window. All measurements can be done on the 2D cross-section or the 3D volume. All traces in the 2D window are shown in their corresponding position on the 3D projection, and vice versa. Measurements include x, y, z coordinates, 2D straight-line distance, angles, perimeters, surface areas, volumes, and true 3D linear distances and surface areas on the 3D projection. In addition, VoxBlast has a lighting model, palette editor, autosegmentation, movie generation and playback, volume subsectioning, 2D image resampling, viewing from any angle, pseudocolouring and transparency.

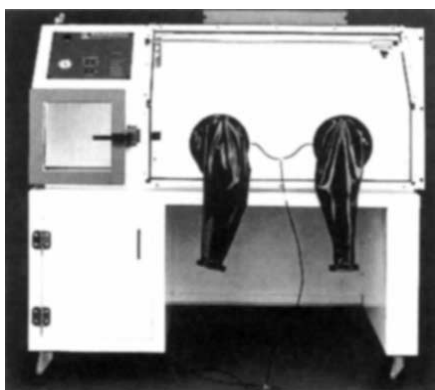
■ Safety cabinets

The main benefit of the new recirculating class I **microbiological safety cabinet** from Envair is the greater flexibility it provides by eliminating the need for ducting to the external atmosphere (*Reader Service No. 113*). In addition, Envair has also redesigned the front of the cabinet, removing the front bar or cross member to enhance the user's visibility of the working area. The company also offers the Bio 2+ class II MSC, which is designed to comply with most clean air standards. Designed as a compact unit to accommodate varied laboratory conditions and limitations, the Bio 2+ offers fast and simple maintenance procedures, with all serviceable items accessible from the front.

The Bactron **controlled environment system** from Sheldon Manufacturing provides 17 cubic foot of controlled atmosphere workspace and gloveless operation for increased productivity and operator comfort (*Reader Service No. 114*). The new system can be used for applications with nitrogen, CO₂, AMG or other inert gases to meet specific requirements. It features airtight sleeves, which seal around the arms of the user, allowing unencumbered, tactile and dextrous gloveless handling of items in the chamber, and permitting items to be brought into the chamber without compromising the



Amray's new field-emission SEM.



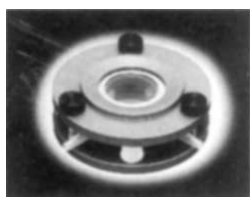
Bactron provides a gloveless workspace in a controlled environment.

integrity of the chamber atmosphere. Work is loaded into and from the chamber through Bactron's air-lock module, which automatically exchanges atmospheres and has a sliding shelf. Stainless steel, airtight construction and a rigid Plexiglas face ensure unobstructed view of work in progress. Other features include microprocessor controls, vacuum pump, microscope adapter, gas regulator and hand-held leak detector.

■ Labware

The new FOTO/UV series of **ultraviolet light transilluminators** from Fotodyne is designed to save on bench space without sacrificing viewing size (*Reader Service No. 115*). The filter glass has been optimized to accommodate popular gel sizes: 15 × 15 cm, 21 × 21 cm and 21 × 26 cm. Both preparative and analytical light settings can be selected by the user. The optional UV blocking shield is hinged for safety and convenience. Fotodyne is also making available a version with stainless steel enclosures for use in corrosive environments.

Medical Systems is distributing a **closed 'micro' perfusion chamber** for single cell fluorescence measurements, which has been designed by the University of Leiden for use on the stage of upright or inverted microscopes (*Reader Service No. 116*). The system consists of standard circular glass coverslips (24/25 mm diameter) that are attached

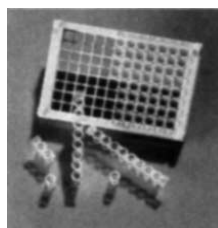


Closed perfusion chamber.

by snap-type fittings. Cells are grown on these coverslips and are attached to the top or bottom of the chamber depending on the type of microscope used. Short optical paths between coverslips (14 mm) allow fluorescence, phase contrast, interference contrast and confocal microscopy. There are eight conically drilled holes in the side of the chamber. These holes can be used to insert perfusion lines, special purpose probes and a thin spiral stainless steel tube to

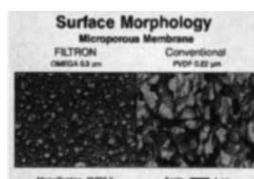
control the chamber environment when perfusion is stopped. Stated uses include measurements of calcium concentration, pH or membrane potential using fluorescent dyes while under physiological conditions.

New **streptavidin-coated microplates, strips and breakable strips** from Lab-systems enable the biotin/streptavidin system to be used for binding DNA, peptides, polysaccharides and other biotin conjugates to the polystyrene well surface (*Reader Service No. 117*). The company says that this system can overcome the binding problems associated with steric hindrance or hydrophilic molecules, and is particularly suitable for solid-phase DNA hybridization assays in addition to immunoassays, receptor studies and other solid-phase assays. Each well of the microplate or strip binds 6×10^{12} molecules of biotinylated oligonucleotide or approximately 10^{12} molecules of 100–500 base-pairs long DNA fragments. In addition to clear plates, Lab-systems manufactures white microplates and strips for those assays utilizing luminescence markers. The products are also available with a black pigment for fluorescence assays. Those researchers using only part of a plate for each assay may be interested in the new breakable streptavidin strips designed to eliminate waste. For assays using radioisotopic markers, the individual wells can be snapped off and placed straight into counting vials.



Streptavidin-coated microplates.

Asymmetric **microfiltration membranes** designed specifically for tangential flow filtration have been developed by Filtron with 0.16 µm and 0.3 µm pore sizes (*Reader Service No. 118*). Unlike conventional microporous membranes that can be prone to irreversible plugging, Filtron says that these membranes, which are suitable for continuous cell processing applications, feature a smooth surface that minimizes fouling from cell entrapment, which is a primary factor of flow rate decay. As a consequence, cleaning of the membrane is simplified and reproducible flow rates after cleaning can be achieved. In addition, the asymmetric interstitial pore structure prevents fouling beneath the surface where further plugging can occur.



Filtron's membranes for cell processing.

Du Pont has introduced two new **superspeed centrifuges**, both of which use

a CFC replacement coolant (*Reader Service No. 119*). The RC-5C Plus and RC-5B Plus floor model centrifuges use the company's new 'Suva' replacement for CFC-12, the conventional chlorofluorocarbon compound used for refrigeration. The new Plus series is also designed to reduce noise levels and heat exhaust. The Superlite SLA 1000 fixed angle rotor is the first of a new family of Superlite aluminium rotors weighing up to 50 per cent less than conventional rotors. In addition, the Sorvall rotor systems offer maximum protection against aerosol spread of infectious, radioactive and other hazardous materials should tubes break or fail. The RC-5C Plus model features an easy-to-use digital control panel that displays temperature, time, speed and diagnostic indicators. Run-to-run reproducibility and automatic g-force determination provide precise, reproducible runs. A built-in automatic rate control offers improved density gradient separations/yield. The RC-5B Plus has an analog format. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more information, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

Micromanipulators

Easy to Use Systems for:

- Microdissection
- Probing
- Soldering

SINGER
INSTRUMENTS

"A responsibility to Science"

Tel: (+44) 984 40226 Fax: (+44) 984 41166

Singer Instrument Co. Ltd., Roodwater, Watchet, Somerset, England, TA23 0QJ, UK.

Reader Service No.50

Teflon[®] Coated and Bare Wire from MEDWIRE

Medwire provides the highest quality platinum, platinum alloy, silver and stainless steel wire in both single and multistrand configurations.

For prompt delivery or additional information, please contact:

MEDWIRE, an affiliate of Sigmund Cohn Corp.
121 So. Columbus Ave. Mount Vernon, NY 10553
Tel: (914) 664-5300 Fax: (914) 664-5377

*Reg TM of E I DuPont & Co

Reader Service No.25

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without interruptions, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to presubmission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are typeset from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose, and to mention availability of these data.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review articles survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

Letters are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and should contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned but suggestions can be made to the Commentary Editor in the form of a one-page synopsis.

News and Views articles inform non-specialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with submitted manuscripts, and clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers are required, together with five copies of a letter detailing the changes made.

Titles are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books should include publisher, place and date.

Abbreviations, symbols, units and greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

Acknowledgements are brief and appear after the reference list; grant and contribution numbers are not allowed.

Supplementary information is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax.